

Carlos Manuel Morais da Costa

Produção e Desempenho Hospitalar Aplicação ao Internamento

**Tese de candidatura ao grau de
Doutor em Saúde Pública
na especialidade de Administração de Saúde pela
Universidade Nova de Lisboa**

**Escola Nacional de Saúde Pública
Universidade Nova de Lisboa
Lisboa
Janeiro 2005**

Carlos Manuel Morais da Costa

**Produção e Desempenho Hospitalar
Aplicação ao Internamento**

Tese de candidatura ao grau de
Doutor em Saúde Pública
na especialidade de Administração de Saúde pela
Universidade Nova de Lisboa

Escola Nacional de Saúde Pública
Universidade Nova de Lisboa
Lisboa
Janeiro 2005

In Memoriam
Augusto Mantas

Dedicatória

À Graça, ao João e ao Pedro
Aos meus pais.

Agradecimentos

Ao longo da realização desta tese, bem como do meu percurso profissional recebi diversos apoios, colaborações e estímulos, dos quais destaco:

Fernando Galvão de Melo, meu amigo e que aceitou co-orientar esta tese. Os seus incentivos, sugestões e “capacidade de síntese” muito contribuíram para a realização deste estudo.

José Pereira Miguel, meu amigo e que aceitou co-orientar esta tese. A sua disponibilidade e a permanente confiança nas minhas capacidades, bem como os seus incentivos e contributos metodológicos contribuíram igualmente de forma decisiva para a realização deste estudo.

Vasco Reis, pela amizade, disponibilidade e entrega, os quais constituíram elementos chave na definição do meu percurso profissional e naturalmente desta tese.

Graça Freitas, pela cumplicidade e espírito de ajuda, sem os quais a realização da tese teria sido mais difícil.

Jorge Simões, Lopes Martins, Rui Janeiro e Silvino Alcaravela, pela amizade e ainda pelo facto de terem sido dos primeiros a acreditar e a “militar” na minha causa da gravidade e do ajustamento pelo risco.

José Gíria e Paulo Ferrinho pela forma desinteressada com que disponibilizaram as bases de dados dos resumos de alta dos hospitais, sem as quais este estudo não poderia ter sido realizado.

Catarina Sena, Paulo Nogueira e Sílvia Lopes, pela total disponibilidade e por todo o apoio logístico facultado para a realização da tese.

Dan Louis e Joe Gonnella, pela amizade, confiança e entusiasmo que sempre me disponibilizaram. “É tão bom e fácil fazer novos amigos”.

Mário Faria, amigo e cúmplice. “Os optimistas cépticos, também têm direito à vida”.

José Nogueira da Rocha, amigo. “Uma história sobre a atracção dos opostos”.

António Correia de Campos, pelo percurso profissional, apesar de todas as divergências.

Francisco Ramos, João Pereira, Manuel Delgado e Rogério de Carvalho, “compagnons de route”, mas não de caminho.

Luís Graça e Carlos Silva Santos, pela amizade, por vezes nos momentos mais difíceis.

Luís Campos, Miguel Gouveia e Vítor Ramos, companheiros insuperáveis de diversas consultorias.

A todos os outros que de diversas formas contribuíram para este ponto de chegada.

e

“Augustinho”, Professor, Conselheiro, Companheiro e Amigo. “Uma história de total cumplicidade”.

Resumo

O presente estudo pretende analisar a produção e o desempenho dos hospitais em Portugal, tendo como limitação o estado da arte dos sistemas de informação existentes.

Neste sentido, após a delimitação do estudo ao âmbito do internamento e a três anos – 1999, 2000 e 2001 – optou-se por utilizar a base de dados dos resumos de alta dos hospitais.

Esta base de dados que foi iniciada em Portugal em 1984, a propósito da introdução dos DRGs como instrumento de financiamento hospitalar, tem vindo a ser aperfeiçoada durante este período, constituindo um dos mais interessantes e poderosos elementos do actual sistema de informação hospitalar.

Atendendo a que uma das principais críticas feitas aos Diagnosis Related Groups (DRGs), tanto para finalidades de financiamento, como para a gestão dos hospitais, tem sido a ausência de mecanismos de medição da gravidade dos doentes, neste estudo apresenta-se como primeira inovação a identificação e caracterização da produção dos hospitais tendo em conta a gravidade dos casos tratados.

Para tal, em função das limitações de dados existentes utilizou-se um sistema de classificação de doentes – o Disease Staging – o qual para além de atribuir estadios de gravidade, está baseado num modelo biológico, no qual as características e os procedimentos dos hospitais são pouco relevantes.

Com a identificação da gravidade dos doentes no panorama hospitalar português foi possível utilizar outro instrumento muito conhecido e vulgarizado no panorama internacional – o ajustamento pelo risco.

Com o ajustamento pelo risco, pretende-se controlar as características dos doentes que podem influenciar ou condicionar os resultados de saúde, procurando-se ainda eliminar ou reduzir os argumentos referidos pelos hospitais e/ou prestadores que defendem que os resultados do meu hospital ou dos meus doentes são diferentes, simplesmente pelo facto de os meus doentes serem igualmente distintos e mais graves.

Tendo em vista o modelo conceptual das dimensões de risco e as características do Disease Staging, utilizou-se o “software” deste programa que disponibiliza diversos “outputs”, dos quais foram somente considerados os estadios e sub-estadios da doença principal e das comorbilidades, bem como as previsões para a mortalidade e para a duração de internamento.

Atendendo a que os dados originais do Disease Staging estão calibrados aos dados de origem (EUA) recorreu-se a um processo de recalibração aos dados portugueses da mortalidade esperada e da duração de internamento esperada.

Para a mortalidade esperada utilizou-se a regressão logística para se recalibrarem os dados, enquanto que para a duração de internamento os valores originais do Disease Staging foram ponderados por valores de referência.

Dado que este processo de recalibração teve sucesso, pode referir-se uma segunda inovação neste estudo, a possibilidade de se terem valores esperados para a mortalidade e para a duração de internamento e concomitantemente poder ajustar pelo risco a actividade dos hospitais.

Tendo em atenção os objectivos e finalidades do estudo, bem como os instrumentos anteriormente referidos foram identificadas as áreas e as dimensões objecto de análise.

Como áreas foram identificadas a produção e o desempenho hospitalar. Na produção hospitalar foram identificadas duas dimensões: a sua medição (índice de casemix) e o perfil das admissões (concentração da produção, complexidade e gravidade). Para o desempenho foram identificadas três dimensões: a efectividade dos cuidados prestados, a eficiência e um modelo global (misto) que inclui as duas dimensões anteriormente referidas.

Nesta conformidade os hospitais foram avaliados em função da sua produção e desempenho. No que se refere à produção, para além da identificação da concentração e da complexidade e da gravidade dos casos tratados por hospital, foram ainda analisadas e discutidas as principais consequências dos perfis produtivos dos hospitais.

Em relação ao desempenho foram introduzidos “rankings” para qualquer uma das dimensões, para o total das admissões, para os casos médicos e cirúrgicos e ainda por doença. Paralelamente foram ainda analisados e discutidos os diferentes desempenhos, tanto no que se refere à efectividade e eficiência dos cuidados prestados, bem como nos diferentes comportamentos internos dos hospitais em função da respectiva perspectiva de avaliação.

Abstract

The main goal of this study is to analyze the Portuguese hospitals production and performance, taking into account the existent health information systems.

Then this study focuses on the inpatient episodes and for the period from 1999 to 2001, hence the choice was made to use an administrative database.

This data was first introduced in 1984 when DRGs began to be used as a tool for the hospital prospective payment system.

Since one of the most common criticism to the DRGs, when they are used for financing as well as for management purposes, is the lack information on severity of illness, this study throughout the utilization of Disease Staging also allows the measurement of severity of illness.

The Disease Staging is a patient classification system that also uses the administrative databases and identifies different stages of severity. Beyond this, it is based on a biological model, for which the characteristics of the hospitals and the procedures performed are almost irrelevant.

The measurement of the severity of illness made possible that a well known tool used in international studies – the risk adjustment – could be introduced in Portugal.

The purpose of risk adjustment is to account for the patients' characteristics that may influence health outcomes and thereby to reduce or even eliminate the arguments that the providers present saying that "my outcomes are different because my patients are different and sicker".

The Disease Staging software was run and some of the several outputs available were used, namely the stages and sub-stages of the principal disease category and of the comorbidities as well as the predicted values for mortality and length of stay (scales).

Because these scales of Disease Staging were calibrated for the USA data, they have been recalibrated for Portuguese data. For the mortality scale the recalibration was made using logistic regression and for the length of stay scale the recalibration process was made throughout a reference value.

The fact that the recalibration process was successful allows this study to use the expected values of mortality and length of stay as means for the risk adjustment.

Taking account these issues several dimensions were considered for measuring the hospital production and to evaluate the hospital performance.

Concerning the hospital production its concentration, complexity and severity of illness were identified. Moreover, the hospital differences were also analyzed and discussed taking into account some institutional characteristics.

The hospital performance evaluation was made for its effectiveness, efficiency and for the mixed approach. Then the hospitals rankings are presented for the overall production and for surgical and medical conditions and also by disease. Finally, these results are discussed specially for the effectiveness and efficiency performance and for the hospital internal differences.

Índice

Prefácio	I
1. Introdução	1
2. Enquadramento Teórico e Questões Práticas	9
2. Enquadramento Teórico e Questões Práticas – Avaliação da Produção e do Desempenho Hospitalar ..	11
2.1. Identificação e Medição da Produção Hospitalar	13
2.2. Perfil de Admissões Hospitalares	17
2.3. Avaliação do Desempenho Hospitalar	21
2.3.1. Importância do Tema	21
2.3.2. Algumas Experiências Internacionais	26
2.3.3. Ajustamento pelo Risco – Conceptualização e Operacionalização	35
2.3.3.1. Razão de Ser	35
2.3.3.2. Dimensões do Risco	37
2.3.3.3. Resultados de Saúde e Efectividade dos Cuidados Prestados	43
2.3.3.4. Resultados de Saúde e Eficiência dos Cuidados Prestados	50
2.3.3.5. Modelos de Ajustamento pelo Risco	56
2.3.3.6. Propriedades dos Modelos de Ajustamento pelo Risco	75
2.3.3.7. Validação dos Modelos de Ajustamento pelo Risco	80
3. Objectivos	85
4. Metodologia	91
4.1. População em Estudo e Critérios de Exclusão	95
4.2. Variáveis em Estudo	97
4.2.1. Definições Conceptuais	97
4.2.2. Definições Operacionais	98
4.3. Instrumentos	107
4.4. Procedimentos para Aplicação do “Software” e Recalibração	111
5. Resultados	139
5.1 Produção Hospitalar	141
5.1.1. Medição da Produção Hospitalar	141
5.1.2. Perfil das Admissões	152
5.2. Desempenho	187
5.2.1. Efectividade	187
5.2.1.1. Caracterização da Taxa de Mortalidade Observada	188
5.2.1.2. Caracterização da Taxa de Mortalidade Esperada	193
5.2.1.3. Avaliação da Efectividade	197
5.2.2. Eficiência	212
5.2.2.1. Demora Média	213
5.2.2.2. Avaliação da Eficiência	216
5.2.3. Desempenho Global	228
5.3. Síntese de Resultados	241
6. Discussão.....	267
6.1. Discussão Metodológica	269
6.1.1. Sector de Avaliação	270
6.1.2. Áreas de Avaliação	272
6.1.3. Dimensões	277

6.1.4. Indicadores e Procedimentos	281
6.1.5. Instrumentos	294
6.1.6. Fontes de Informação	301
6.2. Discussão dos Resultados	305
6.2.1. Produção	305
6.2.1.1. Medição da Produção	305
6.2.1.2. Perfil das Admissões	310
6.2.2. Desempenho	320
6.2.2.1. Efectividade	320
6.2.2.2. Eficiência	341
6.2.2.3. Desempenho Global	359
6.3. Implicações da Avaliação da Produção e do Desempenho no Sector da Saúde	369
7. Conclusões	379
Bibliografia	395

Índice de Quadros

Quadro	Título	Capítulo	Página
I	Reino Unido: Lista de Indicadores do Desempenho Hospitalar	2.3	31
II	Principais Sistemas de Ajustamento pelo Risco	2.3	67
III	Comparação dos Diferentes Sistemas de DRGs	2.3	70
IV	Comparação entre um Modelo com Resultados Dicotômicos e um Modelo com Previsões Dicotômicas	2.3	79
V	População Inicial e População Final; Motivos para a Eliminação de Episódios	4.1	96
VI	Discriminação (estatística “c”) da Taxa de Mortalidade Esperada Por Hospital, após Recalibração	4.2	101
VII	Taxas de Mortalidade Hospitalar Observada e Esperada: 1999, 2000, 2001 e Total do Período	4.4	115
VIII	Taxas de Mortalidade, Observada e Esperada Por Tipo de Hospital e Por Região	4.4	116
IX	Critérios para a Recalibração do Disease Staging: Casos Cirúrgicos	4.4	119
X	Situações sem Calibração: Casos Cirúrgicos	4.4	119
XI	Critérios para a Recalibração do Disease Staging: Casos Médicos	4.4	120
XII	Situações sem Calibração: Casos Médicos	4.4	121
XIII	Ajustamento do Modelo Disease Staging (2º passo): Casos Cirúrgicos	4.4	126
XIV	Critérios para a Recalibração do Disease Staging (2º passo): Casos Cirúrgicos	4.4	126
XV	Ajustamento do Modelo Disease Staging (2º passo): Casos Médicos	4.4	128
XVI	Ajustamento do Modelo Disease Staging: Tumor maligno da mama feminina, Casos Médicos	4.4	129
XVII	Ajustamento do Modelo Disease Staging: Doença Cerebrovascular, Casos Médicos	4.4	129
XVIII	Ajustamento do Modelo Disease Staging: Pneumonia bacteriana, outra, Casos Médicos	4.4	130
XIX	Demora Média Observada e Demora Média Esperada por Doença e para o Total de Doentes/Episódios de Internamento	4.4	133
XX	Índices de Casemix segundo os DRGs e o Disease Staging por Tipo de Hospital e por Região de Saúde	5.1	141
XXI	Índices de Casemix segundo os DRGs e o Disease Staging por gravidade	5.1	147
XXII	Índices de Casemix segundo os DRGs e o Disease Staging por Idade	5.1	147
XXIII	Níveis de Complexidade e Graus de Severidade: Total de Doentes, Casos Cirúrgicos e Casos Médicos	5.1	158
XXIV	% de Doentes Tratados por Níveis de Complexidade e Estádios de Gravidade: Todos os Doentes, Casos Cirúrgicos e Casos	5.1	159

Quadro	Título	Capítulo	Página
	Médicos		
XXV	Produção Média, Complexidade Média, Gravidade Média e Diferenças entre Complexidade e Gravidade por Hospital por Quartis de Produção: Todas as Admissões, Casos Cirúrgicos e Casos Médicos	5.1	179
XXVI	% Doentes Tratados por Estadio de Gravidade e por Nível de Complexidade por Grandes Agrupamentos de Doenças (GADs)	5.1	180
XXVII	Coeficientes de Correlação; Perfil das Admissões Hospitalares	5.1	184
XXVIII	Coeficientes de Correlação por Hospital entre Perfil das Admissões e Resultados Brutos	5.1	185
XXIX	Total de Episódios e Taxa de Mortalidade Hospitalar Observada: 1999, 2000, 2001 e Total do Período	5.2	188
XXX	Taxas de Mortalidade Hospitalar Observadas por Região de Saúde e por Estádios de Gravidade	5.2	188
XXXI	Taxas de Mortalidade Hospitalar Observadas por Tipo de Hospital e por Estádios de Gravidade	5.2	189
XXXII	Hospitais com Taxas de Mortalidade Observadas mais Baixas	5.2	190
XXXIII	Hospitais com Taxas de Mortalidade Observadas mais Elevadas	5.2	191
XXXIV	Taxas de Mortalidade Hospitalar Esperadas por Região de Saúde	5.2	194
XXXV	Taxas de Mortalidade Hospitalar Esperadas por Tipo de Hospital	5.2	194
XXXVI	Hospitais com Taxas de Mortalidade Observadas e Esperadas mais Baixas, por Região e Tipo de Hospital	5.2	196
XXXVII	Hospitais com Taxas de Mortalidade Observadas e Esperadas mais Elevadas, por Região e Tipo de Hospital	5.2	196
XXXVIII	Ordenação dos Hospitais por Efectividade (10 melhores e 10 piores): Casos Cirúrgicos e comparação com o Total de Casos e os Casos Médicos	5.2	205
XXXIX	Ordenação dos Hospitais por Efectividade (10 melhores e 10 piores): Casos Médicos e comparação com o Total de Casos e os Casos Cirúrgicos	5.2	207
XL	Hospitais com Maior e Menor Efectividade por Grandes Agrupamentos de Doenças	5.2	209
XLI	Demoras Médias por Região de Saúde e Tipo de Hospital	5.2	214
XLII	Demoras Médias Observadas e Esperadas por Região de Saúde e por Tipo de Hospital, Razão e “Z score”	5.2	217
XLIII	Concordância na Atribuição de Ordenações por Hospital (K de Cohen) para as Demoras Médias Observadas e Esperadas por Estadio e para o Total dos Doentes	5.2	220
XLIV	Associação entre as Ordenações Globais e por Estádios de Severidade	5.2	221
XLV	Hospitais Mais e Menos Eficientes por Grandes Agrupamentos de Doenças	5.2	226
XLVI	Hospitais com Melhor e Pior Desempenho Global por Grandes Agrupamentos de Doenças	5.2	239
XLVII	10 Hospitais com Valores mais Elevados dos Índices de Casemix Composto, dos DRGs e do Disease Staging; 10 Hospitais com Maiores e Menores diferenças entre os Índices de Casemix dos DRGs e do Disease Staging: Total das Admissões	5.3	241

Quadro	Título	Capítulo	Página
XLVIII	10 Hospitais com Maior Dispersão nas Admissões (DRGs e Disease Staging); 10 Hospitais com maiores e menores diferenças entre a Dispersão: Total das Admissões	5.3	242
XLIX	10 Hospitais com Maior Dispersão nas Admissões a 80% (DRGs e Disease Staging); 10 Hospitais com Maiores e Menores Diferenças entre a Dispersão: Total das Admissões	5.3	243
L	10 Hospitais com Admissões Mais Complexas e Graves; 10 Hospitais com Maiores e Menores Diferenças entre a Complexidade e Gravidade dos Casos Admitidos: Total das Admissões	5.3	244
LI	10 Hospitais com Maior Efectividade: Todas as Admissões, Casos Cirúrgicos e Casos Médicos; 10 Hospitais com Maiores e Menores Diferenças entre os Casos Cirúrgicos e os Casos Médicos	5.3	246
LII	10 Hospitais com Maior Eficiência: Todas as Admissões, Casos Cirúrgicos e Casos Médicos; 10 Hospitais com Maiores e Menores Diferenças entre os Casos Cirúrgicos e os Casos Médicos	5.3	247
LIII	10 Hospitais com Melhor Desempenho Global: Todas as Admissões, Casos Cirúrgicos e Casos Médicos; 10 Hospitais com Maiores e Menores Diferenças entre os Casos Cirúrgicos e os Casos Médicos	5.3	248
LIV	Hospitais com Maiores e Menores Diferenças entre a Efectividade e a Eficiência: Total das Admissões, Casos Cirúrgicos e Casos Médicos	5.3	249
LV	Síntese do Desempenho Hospitalar, Primeira Posição e Cinco Primeiras Posições, Última Posição e Cinco Últimas Posições: Total das Admissões, Casos Cirúrgicos e Médicos e Doenças	5.3	251-266
LVI	Variação na Produção Média, na Complexidade Média, na Gravidade Média e nas Diferenças entre Complexidade e Gravidade por Hospital por Quartis: Todas as Admissões	6.2	315
LVII	Coeficientes de Correlação por Hospital entre Admissões e Resultados Ajustados pelo Risco	6.2	318
LVIII	Efectividade – Ordenação do Desempenho dos Hospitais, por Quartis: Casos Cirúrgicos e Casos Médicos	6.2	331
LIX	Efectividade – Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos por Tipo de Hospital	6.2	332
LX	Efectividade – Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos por Região de Saúde	6.2	333
LXI	Produção Média e Coeficiente de Variação; % de Casos Cirúrgicos, Todos os Hospitais e por Grupos de Variação	6.2	334
LXII	Efectividade Global, 10 Melhores Hospitais, Desempenho para os Casos Cirúrgicos e Médicos e por Doença; Ordenação	6.2	336
LXIII	Efectividade Global, 10 Piores Hospitais, Desempenho para os Casos Cirúrgicos e Médicos e por Doença; Ordenação	6.2	338
LXIV	Hospitais com Maiores Diferenças na Efectividade Global, para Casos Cirúrgicos e Médicos e por Doença	6.2	339
LXV	Eficiência – Ordenação do Desempenho dos Hospitais, por Quartis: Casos Cirúrgicos e Casos Médicos	6.2	350

Quadro	Título	Capítulo	Página
LXVI	Eficiência – Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos por Tipo de Hospital	6.2	351
LXVII	Eficiência – Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos por Região de Saúde	6.2	352
LXVIII	Produção Média e Coeficiente de Variação; % de Casos Cirúrgicos, Todos Hospitais e por Grupos de Variação	6.2	352
LXIX	Eficiência Global, 10 Melhores Hospitais, Desempenho para os Casos Cirúrgicos e Médicos e por Doença; Ordenação	6.2	354
LXX	Eficiência Global, 10 Piores Hospitais, Desempenho para os Casos Cirúrgicos e Médicos e por Doença; Ordenação	6.2	356
LXXI	Hospitais com Maiores Diferenças na Eficiência Global, para Casos Cirúrgicos e Médicos e por Doença	6.2	357
LXXII	Resultados do Modelo; Desempenho Global: Bruto e Ajustado; Total das Admissões, Casos Cirúrgicos e Casos Médicos	6.2	363
LXXIII	Ordenação do Desempenho dos Hospitais, por Quartis: Casos Cirúrgicos e Casos Médicos	6.2	364
LXXIV	Ordenação do Desempenho dos Hospitais, por Quartis, Efectividade e Eficiência	6.2	366
LXXV	Coeficientes de Correlação entre a Efectividade e a Eficiência dos Hospitais	6.2	367
LXXVI	Origem dos Doentes e Prestação de Cuidados, Complexidade e Gravidade dos Casos	6.3	373

Índice de Figuras

Figura	Título	Capítulo	Página
1	Necessidade de Avaliação das Organizações de Saúde (1)	2.3	22
2	Necessidade de Avaliação das Organizações de Saúde (2)	2.3	24
3	Estado de Arte da Avaliação do Desempenho	2.3	27
4	Curva ROC (exemplo demonstrativo)	2.3	80
5	Aspectos Conceptuais e Metodológicos dos Modelos de Ajustamento pelo Risco	2.3	81
6	Hierarquia para a Validação Externa de Sistemas de Ajustamento pelo Risco	2.3	81
7	% de <i>Outliers</i> por Hospital; Reduções na Demora Média por Hospital	4.4	136
8	Demora Média Observada por Hospital; Demora Média Esperada por Hospital (dados originais) (1); Demora Média Esperada por Hospital (após recalibração) (2)	4.4	137
9	Índices de Casemix segundo os DRGs e o Disease Staging por Hospital	5.1	142
10	Índices de Casemix segundo os DRGs e o Disease Staging por Hospital: Casos Cirúrgicos	5.1	144
11	Índices de Casemix segundo os DRGs e o Disease Staging por Hospital: Casos Médicos	5.1	145
12	Índices de Casemix segundo os DRGs e o Disease Staging por Doença	5.1	146
13	Índice de Casemix Composto por Hospital	5.1	150
14	% de Produtos Diferentes por Hospital segundo os DRGs e o Disease Staging	5.1	153
15	% de Produtos Diferentes Responsáveis por 80% da Produção, por Hospital segundo os DRGs e o Disease Staging	5.1	154
16	Concentração da Produção por Hospital: Total de Produtos Diferentes e Produtos Responsáveis por 80% da Produção – DRGs e Disease Staging	5.1	156
17	% de Doentes por Níveis de Complexidade: Total de Admissões	5.1	161
18	% de Doentes por Níveis de Complexidade: Casos Cirúrgicos	5.1	162
19	% de Doentes por Níveis de Complexidade: Casos Médicos	5.1	164
20	Ordenação dos Hospitais por Complexidade (“Z score”): Total de Doentes e Casos Cirúrgicos e Médicos	5.1	165
21	% de Doentes por Estádios de Severidade e Severidade Média por Hospital: Todas as Admissões	5.1	167
22	% de Doentes por Estádios de Severidade e Severidade Média por Hospital: Casos Cirúrgicos	5.1	169
23	% de Doentes por Estádios de Severidade e Severidade Média por Hospital: Casos Médicos	5.1	170
24	Ordenação dos Hospitais por Gravidade (“Z score”): Total de Doentes e Casos Cirúrgicos e Médicos	5.1	172
25	“Z score” das Admissões por Hospital, Complexidade e Gravidade: Todas as Admissões	5.1	174

Figura	Título	Capítulo	Página
26	“Z score” das Admissões por Hospital, Complexidade e Gravidade: Casos Cirúrgicos	5.1	175
27	“Z score” das Admissões por Hospital, Complexidade e Gravidade: Casos Médicos	5.1	176
28	“Z score” das Admissões por Doença, Complexidade e Gravidade: Todas as Admissões	5.1	183
29	Taxas de Mortalidade Hospitalar Observadas por Hospital	5.2	190
30	Taxas de Mortalidade Hospitalar Observadas por Hospital: Total de Admissões (TMT), Casos Cirúrgicos (TMCir) e Casos Médicos (TMMed)	5.2	191
31	Taxas de Mortalidade Hospitalar Observadas por Doença e por Estádios de Gravidade	5.2	192
32	Taxas de Mortalidade Esperadas por Hospital: Todas as Admissões, Casos Cirúrgicos e Casos Médicos	5.2	195
33	Razão entre Taxas de Mortalidade Observadas e Esperadas (TMP) por Hospital	5.2	198
34	Ordenação dos Hospitais por Nível de Efectividade (RTMP), Taxa de Mortalidade Observada (RTM) e Taxa de Mortalidade Esperada (RTME)	5.2	199
35	Ordenação dos Hospitais por Nível de Efectividade, Padronização Directa (“z score”) (RZs) e Indirecta (RTMP)	5.2	200
36	Ordenações pelo Método Indirecto: Total de Casos e Casos Cirúrgicos e Médicos	5.2	202
37	Ordenações pelo Método Directo: Total de Casos e Casos Cirúrgicos e Médicos	5.2	203
38	Ordenações pelos Métodos Directo e Indirecto: Casos Cirúrgicos	5.2	203
39	Ordenações pelos Métodos Directo e Indirecto: Casos Médicos	5.2	204
40	Demora Média por Hospital	5.2	215
41	Demoras Médias Observadas e Esperadas por Hospital: Estadio 1	5.2	217
42	Demoras Médias Observadas e Esperadas por Hospital: Estadio 2	5.2	218
43	Demoras Médias Observadas e Esperadas por Hospital: Estadio 3	5.2	218
44	Demoras Médias Observadas e Esperadas por Hospital: Total de Doentes	5.2	219
45	Ordenações dos Hospitais pelo Método Indirecto: Estádios 1, 2 e 3 e Total de Doentes	5.2	222
46	Ordenações dos Hospitais pelo Método Directo: Estádios 1, 2 e 3 e Total de Doentes	5.2	223
47	Ordenação dos Hospitais, Eficiência: Casos Médicos e Cirúrgicos e Total dos Doentes	5.2	225
48	Desempenho Global dos Hospitais, Ordenações de Efectividade e de Eficiência	5.2	230
49	Desempenho Global dos Hospitais, “Z score” de Efectividade e de Eficiência	5.2	231
50	Desempenho Global dos Hospitais, “Z score” de Efectividade e de Eficiência (RGI) e Ordenações de Efectividade e de Eficiência (RGA)	5.2	232
51	Ordenação dos Hospitais, Global, Efectividade e Eficiência	5.2	234
52	Ordenação dos Hospitais, Desempenho Global, Efectividade e Eficiência: Casos Cirúrgicos	5.2	236
53	Ordenação dos Hospitais, Desempenho Global, Efectividade e	5.2	237

Figura	Título	Capítulo	Página
	Eficiência: Casos Médicos		
54	Dimensões	6.1	277
55	Indicadores	6.1	281
56	Resultados do Modelo: ICMDRGs	6.2	306
57	Resultados do Modelo: ICMDS	6.2	307
58	Resultados do Modelo: ICMCom	6.2	309
59	Resultados do Modelo: Mortalidade; Total das Admissões	6.2	324
60	Resultados do Modelo: Efectividade; Total das Admissões	6.2	324
61	Resultados do Modelo: Mortalidade; Casos Cirúrgicos	6.2	326
62	Resultados do Modelo: Efectividade; Casos Cirúrgicos	6.2	327
63	Resultados do Modelo: Mortalidade; Casos Médicos	6.2	328
64	Resultados do Modelo: Efectividade; Casos Médicos	6.2	328
65	Resultados do Modelo: Demora Média; Total das Admissões	6.2	344
66	Resultados do Modelo: Eficiência; Total das Admissões	6.2	345
67	Resultados do Modelo: Demora Média; Casos Cirúrgicos	6.2	346
68	Resultados do Modelo: Eficiência; Casos Cirúrgicos	6.2	346
69	Resultados do Modelo: Demora Média; Casos Médicos	6.2	348
70	Resultados do Modelo: Eficiência; Casos Médicos	6.2	348

Siglas e Acrónimos

5 Prim	Cinco Primeiros
5 Ult	Cinco Últimos
ACHS	Australian Council on Health Standards
AIM	Acuity Index Method
APACHE	Acute Physiology and Chronic Health Evaluation
APDRGs	All Patient Diagnosis Related Groups
APRDRGs	All Patient Refined Diagnosis Related Groups
CAP	Community-acquired pneumonia
CID-8	Classificação Internacional das Doenças – 8ª Revisão
CID-9-MC	Classificação Internacional das Doenças – 9ª Revisão – Modificações Clínicas
CPHA	Commission on Professional and Hospital Activities
CSI	Computerized Severity Index
CSP	Complications Screening Program
DGS	Direcção Geral da Saúde
DIE	Duração de Internamento Esperada
Dif	Diferença
DIO	Duração de Internamento Observada
DME	Demora Média Esperada
DME1	Demora Média Esperada, estadio 1
DME2	Demora Média Esperada, estadio 2
DME3	Demora Média Esperada, estadio 3
DMET	Demora Média Esperada, total
DMO	Demora Média Observada
DMO1	Demora Média Observada, estadio 1
DMO2	Demora Média Observada, estadio 2
DMO3	Demora Média Observada, estadio 3
DMOT	Demora Média Observada, total
DRGs	Diagnosis Related Groups
DS	Disease Staging
DSdos	Doentes Saídos
DXCAT	Doenças Secundárias (Disease Staging)
EUA	Estados Unidos da América
GADs	Grandes Agrupamentos de Doenças (Disease Staging)
GCDs	Grandes Categorias Diagnósticas (DRGs)
HCFA	Health Care Financing Administration
HCIA	Healthcare Information Agency
ICMCom	Índice de Casemix Composto
ICMDRG	Índice de Casemix dos DRGs
ICMDS	Índice de Casemix do Disease Staging
IGIF	Instituto de Gestão Informática e Financeira da Saúde

INA	Instituto Nacional de Administração
IPOs	Institutos Portugueses de Oncologia
IRDRGs	International Refined Diagnosis Related Groups
Ln	Logaritmo Natural
LnDI	Logaritmo Natural da Duração de Internamento
LnDIEsp	Logaritmo Natural da Duração de Internamento Esperada
LnPR	Logaritmo Natural do Peso Relativo (DRGs)
Localorig	Variável dicotómica para cruzar a localização dos hospitais com a origem dos doentes
MPM	Mortality Probability Models
NDIE	Nova Duração de Internamento Esperada
Nemtot	Mortalidade prevista recalibrada aos dados portugueses
NHPC	National Health Performance Committee
NHS	National Health Service
NORC	National Opinion Research Center
PDXCAT	Doença Principal (Disease Staging)
PMCs	Patient Management Categories
PR1	Peso Relativo, Nível de Complexidade 1
PR2	Peso Relativo, Nível de Complexidade 2
PR3	Peso Relativo, Nível de Complexidade 3
Prim	Primeiro
PSI	Pneumonia Severity-of-Illness
QMAS	Quality Measurement Advisory Service
RACI	Risk-Adjusted Complications Index
RAMI	Risk-Adjusted Mortality Index
RARI	Risk-Adjusted Readmissions Index
RDRGs	Refined Diagnosis Related Groups
REfect	Ordenação Efectividade, Total
REfectC	Ordenação Efectividade, Casos Cirúrgicos
REfectM	Ordenação Efectividade, Casos Médicos
REfic	Ordenação Eficiência, Total
REficC	Ordenação Eficiência, Casos Cirúrgicos
REficM	Ordenação Eficiência, Casos Médicos
RGA	Ordenação Desempenho Global, média das ordenações da efectividade e da eficiência
RGB	Ordenação Desempenho Global, em função das ordenações da efectividade e da eficiência, mas com maior peso da efectividade
RGC	Ordenação Desempenho Global, em função das ordenações da efectividade e da eficiência, mas com maior peso da eficiência
RGI	Ordenação Desempenho Global, média da efectividade e da eficiência
RGIC	Ordenação Desempenho Global, média da efectividade e da eficiência, Casos Cirúrgicos
RGII	Ordenação Desempenho Global, em função da efectividade e da eficiência, mas com maior peso da efectividade
RGIII	Ordenação Desempenho Global, em função da efectividade e da eficiência, mas com maior peso da eficiência
RGIM	Ordenação Desempenho Global, média da efectividade e da eficiência, Casos Médicos

ROC	Receiver Operating Characteristics Curves
ROE1	Ordenação – Observado versus Esperado, Estadio 1
ROE2	Ordenação – Observado versus Esperado, Estadio 2
ROE3	Ordenação – Observado versus Esperado, Estadio 3
ROET	Ordenação – Observado versus Esperado, Total
RTM	Ordenação Taxa de Mortalidade
RTME	Ordenação Taxa de Mortalidade Esperada
RTMP	Ordenação Taxa de Mortalidade Padronizada
RTMPCir	Ordenação Taxa de Mortalidade Padronizada, Casos Cirúrgicos
RTMPMed	Ordenação Taxa de Mortalidade Padronizada, Casos Médicos
RTMPTot	Ordenação Taxa de Mortalidade Padronizada
RZs	Ordenação “Z score”
Rzs1	Ordenação “Z score”, Estadio 1
Rzs2	Ordenação “Z score”, Estadio 2
Rzs3	Ordenação “Z score”, Estadio 3
RZsCir	Ordenação “Z score”, Casos Médicos
RZsMed	Ordenação “Z score”, Casos Cirúrgicos
RzsT	Ordenação “Z score”, Total
RZsTot	Ordenação “Z score”, Total
SAPS	Simplified Acute Physiology Score
SDRGs	Severity Diagnosis Related Groups
SevMed	Gravidade/Severidade Média
SNS	Serviço Nacional de Saúde
ST1	Estadio 1
ST2	Estadio 2
ST3	Estadio 3
SUPPORT	Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatment
TM1	Taxa de Mortalidade, Estadio 1
TM2	Taxa de Mortalidade, Estadio 2
TM3	Taxa de Mortalidade, Estadio 3
TMCir	Taxa de Mortalidade, Casos Cirúrgicos
TMECir	Taxa de Mortalidade Esperada, Casos Cirúrgicos
TMEMed	Taxa de Mortalidade Esperada, Casos Médicos
TMETot	Taxa de Mortalidade Esperada, Total
TMMed	Taxa de Mortalidade, Casos Médicos
TMP	Taxa de Mortalidade Padronizada
TMT	Taxa de Mortalidade, Total
UHDDA	Uniforme Hospital Discharge Data Abstract
VIH	Vírus da Imunodeficiência Humana
ZScComp	“Z score” Complexidade
ZScSev	“Z score” Gravidade/Severidade

PREFÁCIO

Prefácio

As organizações de saúde em geral e os hospitais em particular são frequentemente reconhecidos por terem particularidades e especificidades que conferem uma especial complexidade ao seu processo produtivo e à sua gestão (Jacobs, 1974 e Butler, 1995).

Estas questões estão associadas às condições próprias do mercado onde operam, onde avultam os aspectos relacionados com a assimetria de informação, com a limitação na soberania do consumidor, com a existência de uma integração vertical incompleta, com a procura derivada, com a relação de agência e com a irrelevância do factor preços para racionalizar a tomada de decisões (Evans, 1984).

Por outro lado, avultam ainda alguns aspectos internos, como a existência de uma dupla linha de autoridade, uma decisão sectorizada, o carácter multiproduto da sua actividade e a existência de um critério de sucesso, por vezes formal e pouco explícito (Costa e Reis, 1993) que tornam ainda os hospitais mais diferentes das tradicionais organizações que operam no mercado em geral.

Neste sentido, na literatura hospitalar emergem alguns temas como prioritários, tanto na investigação, como na avaliação do seu funcionamento, nomeadamente os relacionados com a produção, com o financiamento, com a qualidade, com a eficiência e com a avaliação do seu desempenho.

Dentro destas áreas, este estudo vai abordar os aspectos relacionados com a produção e com o desempenho hospitalar.

Mesmo assim, atendendo à vastidão do tema a abordagem será limitada ao Internamento e a algumas dimensões, mais concretamente à medição da produção e ao perfil das admissões, para o primeiro aspecto, e à efectividade e à eficiência no que se refere ao desempenho.

A medição da produção hospitalar assume especial importância, tanto por questões estruturais, associadas à natureza intrinsecamente social e delicada do seu produto, como por questões operacionais, que passam desde o simples conhecimento e comparação da sua actividade, a questões mais técnicas como a avaliação dos cuidados prestados.

Além disso, a medição da produção hospitalar ganha também especial relevo, porque existem várias razões, tanto associadas à procura, como à oferta que dificultam a identificação de produtos e naturalmente a sua comparação (Hornbrook, 1982).

Neste sentido, em função do actual estado da arte da informação hospitalar, serão utilizados dois sistemas de classificação de doentes, os Diagnosis Related Groups (DRGs) e o Disease Staging.

Atendendo às características destes sistemas de classificação de doentes, um medindo a complexidade dos casos tratados, os Diagnosis Related Groups (DRGs) (Fetter et al, 1980 e Mcguire, 1991) e o outro medindo a gravidade, o Disease Staging (Gonnella, Hornbrook e Louis, 1984 e Gonnella, Louis e Gozum, 1999), torna-se importante avaliar a produção hospitalar, em função destas duas dimensões.

A este propósito, convém desde já referir que, a medição da gravidade dos casos tratados nos hospitais constitui uma inovação em Portugal, visto que até ao presente momento não existia nenhum sistema que permitisse a concretização desta finalidade.

É igualmente pertinente referir que, pese embora o facto de o afirmado ser verídico para todos os doentes internados nos hospitais portugueses, existem algumas experiências sectoriais, nomeadamente nas Unidades de Cuidados Intensivos, na qual a gravidade é medida e utilizada como auxiliar da decisão terapêutica (Carneiro, 1994).

Ainda dentro da área da produção, é igualmente interessante avaliar outra dimensão – o perfil das admissões hospitalares.

Tal pode ser justificado pelas características do Serviço Nacional de Saúde (SNS) português, onde se devem evidenciar a universalidade dos cuidados e a quase inexistência de barreiras à acessibilidade às organizações de saúde.

Ou ainda, por razões mais pragmáticas, visto que se admite que as características dos doentes, desde a própria doença, até à respectiva gravidade, podem condicionar os resultados em saúde (Hornbrook e Goldfarb, 1983; Gonnella, Hornbrook e Louis, 1984; Schumacher et al, 1987; Thomas e Asfcraft, 1989 e Barnum, Kutzin e Saxenian, 1995).

Assim, na literatura norte-americana, por exemplo, é pacificamente aceite que estas mesmas características dos doentes, por terem implicações nos resultados de saúde e nos resultados financeiros, podem conduzir a mecanismos designados por selecção adversa (Scitovsky, McNall e Benham, 1978; Ellis, 1998; Robinson e Gardener, 1995; Marquis e Buchanan, 1999; Shen e Ellis,

2002; Altman, Cutler e Zechauser, 2003; Meenan et al, 2003; Cao e McGuire, 2003 e Shapiro et al, 2003).

Pelo referido anteriormente em relação ao SNS, este mecanismo poderá estar ausente, mas tal não implica que a análise de eventuais comportamentos diferentes por parte dos hospitais não seja relevante isoladamente, tornando-se ainda mais importante pelas repercussões que poderá originar no desempenho destas organizações de saúde.

O conhecimento e eventual publicitação do desempenho dos hospitais é importante, não só para se poderem comparar estas organizações de saúde, mas igualmente para se poder reduzir a actual assimetria de informação (Griffith, Alexander e Jelinek, 2002).

Assim, a avaliação da efectividade e da eficiência dos hospitais constituem-se como dimensões relevantes para uma perspectiva mais ampla – a avaliação do desempenho dos hospitais (Iezzoni, 1995, Hughes et al, 1996 e Healthcare Information Agency (HCIA), 1999).

A dificuldade em avaliar o desempenho das organizações de saúde tem sido evidenciada em diversos estudos nacionais e internacionais, sendo de referir os aspectos associados com a multiplicidade de perspectivas e de agentes, a existência de uma dupla linha de autoridade e a especificidade na garantia da qualidade dos cuidados prestados (Jacobs, 1974 e Donabedian, 1980).

No entanto, apesar das dificuldades conceptuais e operacionais, torna-se cada vez mais necessário desenvolver modelos para a sua concretização. Esta afirmação é válida para qualquer dos agentes associados ao mercado da saúde.

Para os consumidores, visto que é cada vez mais imperiosa a disponibilização de informação de índole científica e credível que lhes permita conhecer a actividade das organizações de saúde, para que desta forma possam exercer com maior fundamentação a sua liberdade de escolha.

Para os proprietários, já que somente a explicitação de modelos válidos e “universais” lhes permitirá otimizar a sua função de utilidade.

Para os gestores e outros profissionais de saúde, porque a existência de um modelo de avaliação do desempenho das organizações de saúde, permite igualmente a sua extensão para o plano de avaliação individual e desta forma explicitar mecanismos de avaliação da competência e a possível atribuição de incentivos.

Para além destes aspectos sectoriais existe igualmente uma variedade de razões que tornam igualmente imperiosa a definição e implementação de modelos de avaliação do desempenho das organizações de saúde:

- Em primeiro lugar, devem ser referidos os aspectos relacionados com a acessibilidade, podendo este aspecto assumir diversas perspectivas, em função do enquadramento organizacional do sistema de saúde. Em Portugal, para além de imperativo constitucional, este aspecto assume importância de dupla natureza, tanto associado a eventuais políticas de selecção/desnatação praticadas por cada hospital (acessibilidade passiva), como no que se refere à disponibilização de informação que permita aos consumidores uma escolha fundamentada dos prestadores (acessibilidade activa).
- Em segundo lugar, devem ser consideradas as questões associadas ao financiamento das organizações de saúde. A este propósito deve referir-se que independentemente da forma como seja considerada a sua prospectividade – orçamento global, capitação ou pagamento pela produção – somente a definição prévia de mecanismos de avaliação do desempenho poderá de facto criar mecanismos para otimizar o “valor do dinheiro”.
- Finalmente, em termos de gestão das organizações de saúde. De facto, somente após a definição, explicitação e implementação de um modelo de avaliação do desempenho das organizações de saúde, se poderão criar mecanismos que, simultaneamente, proporcionem uma maior responsabilização dos agentes internos (gestores e profissionais prestadores) e possibilitem ainda a sua reprodutibilidade na actividade e na gestão de cada organização de saúde.

Para além disso, existem em Portugal alguns aspectos conjunturais que tornam igualmente imperiosa a explicitação de modelos de avaliação do desempenho.

- Um primeiro grupo de razões está associado aos projectos de reforma existentes. De facto, tanto a criação de um sector público empresarial na saúde, como de parcerias público-privado para a construção de novas unidades, constituem aspecto críticos, tanto para a oferta - somente após a definição de um cenário de avaliação se podem estabelecer modelos previsionais para análises de rentabilidade dos capitais e recursos – como para o estado regulador – somente a criação destes mecanismos de avaliação lhe poderá permitir uma avaliação científica, unívoca e criteriosa da bondade das novas experiências de gestão.
- Um segundo grupo de razões poderá ser encontrado na sociedade portuguesa:
 - Por um lado, em face da exposição que todos sofrem por parte da comunicação social. Neste sentido, pensa-se que a criação

e divulgação de modelos de avaliação do desempenho permitirá uma política mais proactiva e poderá minimizar a proliferação de imagens negativas e parcelares de que actualmente são alvo as organizações de saúde;

- Por outro lado, pela comparação com outros sectores de actividade. É praticamente generalizado o “furor” avaliativo e muitas vezes classificativo existente, sendo um bom exemplo disso a publicação dos “rankings” das escolas do ensino secundário e agora, igualmente na saúde, designadamente nos hospitais SA.

Tendo em vista os aspectos enunciados, qual deverá a ser o caminho a percorrer em Portugal?

Na grande maioria das organizações o seu desempenho é avaliado pelos resultados (Costa e Reis, 1993). Nas organizações não lucrativas em geral e nas organizações de saúde este axioma tem sido posto em causa (Evans, 1981 e McGuire, Henderson e Money, 1988).

Mas, será que este aspecto deverá ser considerado estrutural (ponto de partida) ou, antes pelo contrário, deverá ser perspectivado como uma consequência das dificuldades operacionais e conjunturais das organizações de saúde (ponto de chegada)?

A resposta a esta questão não é fácil, nem imediata.

Até ao presente momento em Portugal a abordagem dos resultados para a avaliação do desempenho das organizações de saúde não tem sido privilegiada, como é patente em alguns estudos realizados, quer pelo Instituto Nacional de Administração (INA) para a comparação de hospitais (INA – “Avaliação dos hospitais Fernando Fonseca e Garcia da Orta”, 1999 e “Projecto de Avaliação de Unidades de Saúde”, 2001), como pelo Instituto de Gestão Informática e Financeira da Saúde (IGIF), como por exemplo no estudo “graus de pertença” (Barros, 2001a; Barros 2001b e Barros, 2001c).

Mais recentemente existe a experiência de avaliação do desempenho e de definição de “rankings” para os hospitais SA. Esta metodologia que pretende “aumentar a performance económico-financeira dos hospitais num contexto de melhoria simultânea da qualidade do serviço prestado aos utentes, nomeadamente a nível do acesso”, inclui alguns indicadores de resultados, essencialmente associados à eficiência – custos e demora média (Nóbrega, 2004 e Ribeiro, 2004).

O estado da arte da avaliação do desempenho das organizações de saúde parece seguir a trilogia definida por Donabedian (1985) – Estrutura, Processo e Resultados.

Em que estrutura pode ser definida pelas características da oferta (organizações de saúde); processo por aquilo que é feito ao doente e resultados pela forma como o doente responde aos cuidados que lhe são prestados (QMAS, 1997).

Na realidade embora esta trilogia tivesse originalmente sido concebida para aplicação na garantia da qualidade dos cuidados prestados, parece igualmente legítimo afirmar que esta perspectiva é aplicável a outras realidades da actividade hospitalar, designadamente da eficiência.

Nas experiências dos EUA, por exemplo, deve referir-se que enquanto na metodologia “Best Hospitals” (Hill, Winfrey e Rudolph, 1997 e O’Muircheartaigh, Murphy e Moore, 2002) e no modelo preconizado por DesHarnais e outros (1997 e 2000) se privilegia essencialmente a dimensão qualidade (com a particularidade de a primeira, encerrar as três dimensões da qualidade, enquanto que a última foca essencialmente os resultados), o modelo desenvolvido pela HCIA (HCIA, 1999 e Griffith, Alexander e Jelinek, 2002), pelo contrário, envolve as perspectivas da qualidade, eficiência e rendibilidade dos capitais, embora considere somente uma dimensão – os resultados.

No Reino Unido, embora existam basicamente duas grandes abordagens, a primeira baseada na definição de indicadores do Serviço Nacional de Saúde (Chang et al, 2002) e a segunda interessada na concretização de uma metodologia para estabelecer “hospital ratings” (NHS, 2002 e Snelling, 2003), ambas englobam indicadores de estrutura, processo e resultados.

Na Austrália, embora os modelos utilizados sejam completamente diferentes, oscilando entre indicadores de qualidade, com primazia no processo, para outros mais gerais, com indicadores sobre financiamento, acesso, eficiência e qualidade, nos quais a perspectiva privilegiada é o processo (Ibrahim et al, 1998; Degeling et al, 2000; NHPC, 2001 e ACHS, 2002).

Esta discussão internacional, para além de não se revelar pacífica, reflecte na generalidade uma filosofia básica: enquanto que nos EUA se pensa que se os resultados das organizações de saúde forem razoáveis então é natural que a estrutura e o processo apresentem o mesmo sentido, no Reino Unido e na Austrália, bem como em outros países europeus, pensa-se que se a estrutura e o processo apresentarem bons indicadores, então é natural esperar que os resultados das organizações de saúde sejam igualmente bons.

Atendendo ao referido anteriormente, deve ser adoptada uma abordagem mista, a qual deve incidir nas áreas da eficiência e da qualidade dos cuidados prestados e incluir as perspectivas da estrutura, do processo e dos resultados.

Contudo, problemas essencialmente de ordem operacional – facilidade e economia de implementação da metodologia de avaliação do desempenho dos hospitais e disponibilidade de dados existentes nos hospitais portugueses – exigem uma definição de prioridades para se atingir este objectivo.

Assim, neste estudo será privilegiada a abordagem de avaliação pelos resultados, tendo essencialmente em conta que somente após terem sido identificadas, tanto as organizações de saúde em melhor situação, como aquelas que apresentem piores indicadores, se devem procurar as razões explicativas para tal desempenho. Estas devem então ser procuradas nas questões associadas à estrutura e ao processo de tratamento dos hospitais.

Para um segundo momento poder-se-á então discutir o modelo global de avaliação do desempenho dos hospitais, no qual devem ser avaliadas as três perspectivas referidas – estrutura, processo e resultados.

No presente momento e atendendo ao estado da arte da avaliação do desempenho das organizações de saúde e ao sistema de informação hospitalar existente em Portugal, o estudo, como foi referido, irá incidir exclusivamente sobre uma componente da produção hospitalar, o Internamento e será centrado nos seguintes indicadores de resultados:

- Efectividade – medida pela relação entre a Taxa de Mortalidade Observada e a Taxa de Mortalidade Esperada
- Eficiência – medida pela relação entre Demora Média Observada e Demora Média Esperada
- Global – construído em função da efectividade e da eficiência

Estas perspectivas são analisadas em três anos – 1999, 2000 e 2001. Atendendo à diversidade existente o estudo será feito para todos os episódio de internamento, para os casos cirúrgicos, para os casos médicos e por doença.

INTRODUÇÃO

1. Introdução

Este trabalho incidirá sobre a produção e o desempenho hospitalar, tendo em vista um diagnóstico da situação e a contribuição com novas perspectivas para a melhoria da prestação e da administração dos hospitais.

Atendendo, por um lado, aos assuntos abordados no Prefácio, onde se releva a importância do problema, bem como são referidos os principais contornos da temática em estudo e ainda a que no capítulo seguinte – Enquadramento Teórico e Questões Práticas – será feita uma revisão do actual estado da arte, na introdução far-se-á somente um enquadramento geral do estudo.

Assim, a avaliação da produção e do desempenho hospitalar constituem os objectos primordiais deste estudo, pelo que a estrutura do trabalho é a seguinte:

- Enquadramento Teórico e Questões Práticas;
- Objectivos;
- Metodologia;
- Resultados;
- Discussão;
- Conclusões.

Por outro lado, deve ainda ter-se presente que nenhum dos capítulos do estudo é perfeitamente estanque, pelo que alguns aspectos serão abordados em diferentes capítulos, o que releva a transversalidade deste trabalho. A título perfeitamente exemplificativo, deve referir-se que embora exista um capítulo específico para a discussão metodológica, a mesma refere-se somente aos aspectos mais gerais, pelo que em cada um dos restantes capítulos da discussão serão abordadas questões metodológicas mais específicas.

No Enquadramento Teórico e Questões Práticas foram consideradas três grandes perspectivas: a metodologia para a medição da produção hospitalar, o perfil das admissões (caracterização da produção) e a avaliação do desempenho hospitalar.

Nos aspectos relacionados com a identificação e com a medição da produção hospitalar, para além de se discutirem quais as preferências que contam e a sua relação com o produto hospitalar, define-se o conceito de produto hospitalar, pelo que em termos fotográficos o caso tratado (episódio de internamento) corresponde ao produto hospitalar.

Em seguida, identificam-se e descrevem-se as abordagens para identificar e comparar casos – os sistemas de classificação de doentes e os índices escalares (ou Índices de Casemix).

Enquanto que os sistemas de classificação de doentes estão mais preocupados com as características dos doentes que permitem estabelecer afinidades susceptíveis de criarem grupos de casos (produtos), os índices escalares tratam dos aspectos que permitem a comparação dos produtos previamente definidos.

No perfil das admissões são discutidos dois tipos de questões: a concentração/diversificação da produção e a natureza dos casos tratados, especialmente a sua complexidade e a sua gravidade.

Com a concentração/diversificação pretende-se estudar a proporção de produtos diferentes tratados em cada hospital e consequentemente a raridade dos casos tratados, tendo essencialmente em atenção se a raridade está ou não a ser confundida com a complexidade e/ou com a gravidade.

A complexidade é a medida que expressa a quantidade de recursos necessária para tratar cada caso, portanto um indicador predominantemente da oferta de cuidados de saúde e a gravidade é a medida que expressa a probabilidade de morte ou de falência de um órgão, consequentemente um indicador da procura de cuidados de saúde.

A análise dos comportamentos dos hospitais no que se refere ao seu perfil de admissões é então discutida face aos princípios de universalidade e de acessibilidade associadas a estruturas do tipo Serviço Nacional de Saúde, como é o caso português e ainda sobre a eventual existência de políticas de selecção adversa, muito mais prováveis em estruturas de saúde mais liberais, como é o caso dos EUA.

No que se refere à avaliação do desempenho hospitalar são abordadas questões relacionadas com a sua importância específica, com a descrição de algumas experiências internacionais e com a conceptualização e operacionalização do ajustamento pelo risco.

Na importância da avaliação do desempenho hospitalar, para além de serem identificadas as vantagens e preferências dos diversos agentes que operam no mercado da saúde, foram delineados os aspectos mais relevantes na avaliação do desempenho hospitalar.

Em seguida descrevem-se algumas experiências internacionais, nomeadamente dos EUA, dada a sua grande experiência nesta temática e do Reino Unido, pela sua maior proximidade à lógica organizacional existente em Portugal.

A síntese destas experiências permitiu concluir que a abordagem pelos resultados e dentro destes a efectividade e a eficiência constituem dimensões relevantes para a avaliação do desempenho hospitalar. Foi igualmente referido que para se eliminarem ou, pelo menos controlarem, as características dos doentes que podem originar resultados diferentes é necessário que se realizem ajustamentos pelo risco.

O ajustamento pelo risco, o que é sinónimo de ajustar pela gravidade, é inicialmente contextualizado, através da identificação e caracterização das dimensões do risco.

Em seguida, analisa-se a relação entre os resultados de saúde ajustados pelo risco e a efectividade dos cuidados prestados (resultados ou consequência de determinado procedimento ou tecnologia médica quando aplicados na prática). É igualmente referido que, embora existam opiniões contraditórias, é reconhecida internacionalmente a relação entre aqueles dois aspectos.

À semelhança do efectuado para a efectividade, analisa-se a relação entre resultados ajustados pelo risco e a eficiência (medida que exprime a relação entre recursos utilizados e resultados obtidos). Discutem-se ainda as vantagens e inconvenientes de utilizar dois indicadores – os custos médios ou a demora média – tendo-se evidenciado que na prática existe uma enorme associação entre os mesmos.

São igualmente discutidos os modelos de ajustamento pelo risco, essencialmente as diferenças entre modelos clínicos e administrativos, referindo-se ainda que, embora os primeiros apresentem maior segurança e fiabilidade, existe um grande predomínio nos segundos, principalmente por maior facilidade de implementação e de exploração, bem como da minimização dos custos que estão associados a estas questões.

São ainda abordadas as questões das propriedades estatísticas dos modelos de ajustamento pelo risco, as quais, ao privilegiarem a regressão logística como método preferencial, elegem a calibração e a discriminação como características mais relevantes para se apurar o ajustamento dos modelos.

Com a calibração analisam-se os desvios entre a mortalidade observada e a mortalidade esperada e na discriminação analisa-se se o sistema está a prever taxas de mortalidade mais elevadas para os doentes que efectivamente morreram em relação aos que efectivamente sobreviveram.

A conjugação destes dois atributos permite afirmar que a calibração é mais importante quando se pretendem comparar valores observados com valores esperados e que a discriminação pretende essencialmente diferenciar o risco de morte.

Finalmente, atendendo a que muitas vezes os valores esperados são derivados de outras populações, tanto em termos históricos, como geográficos, é analisada e discutida a necessidade de se recalibrarem esses dados à nova população, referindo-se que, teoricamente a recalibração é sempre um fenómeno único e singular, pelo que novos procedimentos devem ser aplicados quando se estudam novas populações.

Em seguida, são definidos os objectivos do estudo, os quais estão essencialmente relacionados com a aplicação da gravidade e a sua importância para a avaliação da produção e do desempenho hospitalar.

Na Metodologia são considerados quatro aspectos: a população em estudo e os critérios de exclusão, as variáveis utilizadas, os instrumentos utilizados e os procedimentos utilizados para a identificação da gravidade e para a recalibração dos dados.

Em termos gerais foram utilizados os dados dos hospitais públicos em Portugal (com excepção dos Açores e da Madeira, visto que não utilizam a base de dados usada neste estudo), para os anos de 1999, 2000 e 2001, com as grandes exclusões verificando-se nos doentes de neonatologia e nos doentes sem episódios de internamento completos (transferências para outros hospitais, alta contra parecer médico e seguimento em serviço domiciliário).

As variáveis utilizadas, são definidas conceptual e operacionalmente, sendo genericamente as seguintes:

- Produção
 - Medição da Produção
 - Índice de Casemix (Índice de Casemix dos DRGs, Índice de Casemix do Disease Staging e Índice de Casemix Conjunto)
 - Perfil das Admissões
 - Concentração/diversificação total da população (raridade)
 - Concentração/diversificação da produção a 80%
 - Complexidade dos Casos
 - Gravidade dos Casos
 - Complexidade e Gravidade dos Casos
- Desempenho
 - Efectividade
 - Relação entre a taxa de mortalidade observada e a taxa de mortalidade esperada
 - Eficiência
 - Relação entre a demora média observada e a demora média esperada
 - Desempenho global
 - Média não ponderada da efectividade e da eficiência

Os instrumentos utilizados neste estudo respeitam principalmente a dois sistemas de classificação de doentes – os Diagnosis Related Groups (DRGs) e o Disease Staging – os quais são genericamente caracterizados.

Os procedimentos utilizados no estudo respeitam essencialmente ao seguinte:

- Uniformização das versões dos DRGs, visto que no período em análise existiram duas versões dos DRGs;
- Aplicação do “software” do Disease Staging, para identificar a doença principal, os estadios e os sub-estadios da doença principal e das comorbilidades e as previsões para a mortalidade e para a duração de internamento;
- Recalibração da mortalidade esperada;
- Recalibração da duração de internamento esperada.

Os resultados são apresentados para as seguintes dimensões:

- Índice de Casemix calculado pelos DRGs, pelo Disease Staging e Composto, para o total dos episódios de internamento, para os casos cirúrgicos e para os casos médicos;
- Concentração/Diversificação da produção total (efeito raridade);
- Concentração/Diversificação da produção a 80%;
- Complexidade dos casos tratados, para o total dos episódios de internamento, para os casos cirúrgicos e para os casos médicos;
- Gravidade dos casos tratados, para o total dos episódios de internamento, para os casos cirúrgicos e para os casos médicos;
- Efectividade dos hospitais, para o total dos episódios de internamento, para os casos cirúrgicos e para os casos médicos e por doença;
- Eficiência dos hospitais, para o total dos episódios de internamento, para os casos cirúrgicos e para os casos médicos e por doença;
- Desempenho Global dos hospitais (efectividade e eficiência), para o total dos episódios de internamento, para os casos cirúrgicos e para os casos médicos e por doença;
- Indicadores de síntese da produção e do desempenho dos hospitais.

A discussão é feita em relação aos seguintes aspectos:

- Discussão Metodológica;
- Discussão da Produção (medição da produção e perfil das admissões);
- Discussão do Desempenho (efectividade, eficiência, efectividade versus eficiência e desempenho global);
- Discussão sobre as Implicações no Sector da Saúde (financiamento, função regulação, gestão dos hospitais, publicitação dos resultados).

Finalmente são retiradas as conclusões, evidenciando-se as inovações deste estudo e com formulação de algumas recomendações sobre a organização e funcionamento dos hospitais em Portugal, bem como sobre perspectivas de investigação futuras.

ENQUADRAMENTO TEÓRICO E QUESTÕES PRÁTICAS

- Identificação e Medição da Produção Hospitalar
- Perfil de Admissões Hospitalares
- Avaliação do Desempenho Hospitalar

2. Enquadramento Teórico e Questões Práticas

Avaliação da Produção e do Desempenho Hospitalar

Os hospitais são frequentemente reconhecidos como uma das organizações com maior grau de complexidade na sua estrutura e administração. Para tal contribuem uma dinâmica própria e especial do mercado onde operam e a existência de características específicas dos seus modelos económicos e de gestão (Jacobs, 1974 e Evans, 1984).

Entre estas avulta o carácter multiproduto da sua actividade, decorrente de uma enorme diversidade nos diagnósticos/doenças que podem ser tratados no hospital, a qual, por sua vez, pode ainda ser potenciada pelo diferente grau da evolução da doença presente no momento de contacto com o hospital (Hornbrook, 1982 e Tatchell, 1983).

Por outro lado, observa-se ainda um intenso debate sobre temas relacionados com a gestão hospitalar, de entre os quais emerge a qualidade dos cuidados prestados, a eficiência da prestação de cuidados e o financiamento das organizações de saúde (Donabedian, 1985; Costa, 1990 e Butler, 1995).

No entanto, estas questões decorrem do conhecimento e da discussão sobre uma questão central – a definição e a medição da produção hospitalar.

Neste sentido, facilmente se pode perspectivar a decisiva importância que assume o conhecimento e a medição da produção hospitalar, essencialmente no que se refere (Bentley e Butler, 1981; Hornbrook, 1982; Plomman, 1982 e Aronow, 1988):

- Ao conhecimento do que se produz, tanto em termos quantitativos, como qualitativos. Ou seja, refere-se à definição da informação que permita identificar o tipo de doentes tratados em cada organização de saúde. Este aspecto é essencial para a caracterização dos doentes tratados por hospital ou por serviço ou por médico, com as naturais consequências ao nível da gestão e na criação de instrumentos que permitam compreender e justificar o seu desempenho;
- Ao conhecimento de como e quem produz. Esta perspectiva situa-se ao nível técnico, cujos aspectos mais relevantes respeitam à identificação, avaliação e correcção de eventuais anomalias do processo produtivo. Na realidade, esta questão está directamente relacionada com a garantia da qualidade dos cuidados prestados, com a revisão de utilização e com a gestão das situações em risco. Considerando para o efeito a garantia de qualidade como uma abordagem em sede da efectividade, ou seja, com a obtenção de resultados de saúde desejáveis face a “standards” pré-estabelecidos. Por sua vez, a revisão de utilização está associada à adequação e à eficiência dos cuidados prestados, ou seja com

optimização dos recursos para se atingirem os resultados desejáveis. Finalmente, a gestão das situações em risco está relacionada com a identificação e controlo de situações inesperadas em termos dos resultados de saúde;

- Ao apuramento dos custos de produção como elemento facilitador das análises de eficiência e do estabelecimento de um sistema de financiamento com incentivos adequados;
- À identificação de elementos que permitam o conhecimento do funcionamento e do desempenho dos hospitais, com as naturais consequências para todos os agentes que operam no mercado da saúde.

Tendo em atenção os aspectos enunciados as questões mais relevantes a analisar são as seguintes:

- A caracterização da metodologia para se identificar e medir a produção hospitalar;
- A identificação das dimensões relevantes para se caracterizar a produção hospitalar, designadamente os aspectos associados com o perfil das admissões;
- A conceptualização e operacionalização de modelos de avaliação do desempenho hospitalar.

A este propósito convém referir que, atendendo à dimensão e complexidade dos assuntos em análise, serão predominantemente discutidos os aspectos associados com o internamento hospitalar.

2.1. Identificação e Medição da Produção Hospitalar

A definição do produto hospitalar exige a discussão de três questões (Hornbrook, 1982):

- Quais as preferências que contam?
- Qual é a organização relevante?
- Qual é natureza deste produto?

Segundo Berki (1972) o produto hospitalar pode ser abordado de duas formas: (1) o produto hospitalar é o resultado do processo de tratamento, ou (2) o produto é o objecto de troca entre a organização e o consumidor, tendo em vista a optimização da satisfação deste último agente.

Entendendo a organização como o local onde se estabelecem decisões em função das actividades produtivas (Friedman, 1962 e Lipsey e Steiner, 1969, citados em Hornbrook, 1982), o hospital pode ser encarado de duas formas (Hornbrook, 1982):

- Como uma organização constituída por uma variedade de departamentos que produzem produtos distintos (hoteleiros ou de meios complementares de diagnóstico e terapêutica, por exemplo) para os responsáveis pelo tratamento (médicos);
- Como uma organização em que os proprietários, administradores e prestadores são conjuntamente responsáveis pela produção de cuidados de saúde.

Na primeira situação o produto é o serviço, enquanto que na segunda alternativa o caso tratado (episódio) é o produto do hospital.

A conjugação destas questões permite identificar o caso tratado (episódio) como o produto relevante do hospital, designadamente em função das preferências dos consumidores, os quais privilegiam o resultado final do tratamento, em detrimento da quantidade dos produtos intermédios (meios complementares de diagnóstico e terapêutica, por exemplo) que lhe são disponibilizados (Hornbrook, 1982).

Neste sentido, ainda segundo este autor, a natureza do produto respeita à provisão de serviços adequados para melhorar a situação de saúde de cada consumidor.

Neste sentido, os aspectos mais relevantes respeitam à forma de definir, identificar e comparar os casos. Para tal existem genericamente duas abordagens (Hornbrook, 1982):

- Os Sistemas de Classificação de Doentes;
- Os Índices Escalares.

A utilização de sistemas de classificação de doentes pressupõe a definição de critérios que permitam agregar os doentes tratados em cada hospital, enquanto que com os índices escalares (Índices de Casemix) se procura sintetizar num único indicador a actividade do hospital (Horn e Schumaker, 1979; Hornbrook, 1982; Farley e Hogan, 1990 e McGuire, 1991).

Para se desenvolverem sistemas de classificação de doentes os critérios para se identificarem produtos podem respeitar aos sintomas, aos diagnósticos principais, à doença (diagnósticos principais e secundários), aos recursos necessários, aos resultados de tratamento e ao valor social do produto (Hornbrook, 1982).

A aplicação desta metodologia originou uma enorme variedade de sistemas de classificação de doentes como os Diagnosis Related Groups (DRGs), Patient Management Categories (PMCs), MedisGroups, Disease Staging, Computerized Severity of Illness (CSI), Acute Physiology and Chronic Health Evaluation (APACHE), Simplified Acute Physiology Score (SAPS) e Mortality Probability Models (MPM).

Atendendo a que estes sistemas de classificação de doentes se encontram exaustivamente documentados e comparados em diversos estudos (Fetter et al, 1980; Young, Swinkola e Hutton, 1980; Hornbrook, 1982; Brewster, Jacobs e Bradbury, 1984; Gonnella, Hornbrook e Louis, 1984; Young, 1984; Hornbrook, 1985; Knaus et al, 1985; Thomas, Ashcraft e Zimmerman, 1986; Aronow, 1988; Charbonneau et al, 1988; Rosko, 1988; Vladeck e Kramer, 1988; Thomas e Ashcraft, 1991; Costa, 1991; lezzoni et al, 1992a; Markson et al, 1991; lezzoni e Daley, 1992; Le Gall, Lemeshow e Saulnier, 1993; Lemeshow et al, 1993; Costa, 1994; lezzoni, 1997a), serão somente analisados os principais aspectos críticos desta abordagem.

Neste sentido são identificados os seguintes potenciais problemas (Hornbrook, 1982; lezonni et al, 1997):

- O desempenho de cada sistema de classificação é melhor para a finalidade para que foi construído. Tal facto, poderá conduzir à sua proliferação, o que para além de aumentar os custos de implementação e de exploração, pode criar a existência de informações contraditórias sobre a actividade dos hospitais;
- A obrigatoriedade de um doente somente ser classificado numa categoria (produto) pode conduzir a uma maior arbitrariedade na respectiva classificação e eventualmente contribuir para uma maior heterogeneidade de casos dentro do mesmo produto;

- A existência de categorias residuais, como casos não específicos ou outras situações, o que pode constituir um problema essencialmente se nestas categorias forem encontradas frequências elevadas;
- A insuficiência de informação para classificar correctamente um episódio de internamento, a qual pode decorrer de preenchimentos incompletos dos processos clínicos ou ainda de impossibilidade de se conhecer com precisão a situação específica, como por exemplo se pode observar num óbito que ocorre num momento muito próximo da admissão do doente.

Por outro lado, como referido, os índices escalares (Índices de Casemix) são uma medida que pretende sintetizar num único indicador toda a diversidade dos produtos tratados em cada hospital (Horn e Schumaker, 1979; Hornbrook, 1982; Lichtig, 1986).

Basicamente existem duas grandes alternativas para se calcularem índices escalares: a comparação entre a diversidade de casos tratados num hospital em relação à média dos casos tratados num determinado conjunto de hospitais; ou a ponderação dos casos tratados nos hospitais em função de um determinado peso específico (Lichtig, 1986; Farley e Hogan, 1990 e McGuire, 1991).

Para além disso um índice escalar deve conter três componentes (Klasterin e Watts, 1982:

- Um sistema de classificação de doentes que permita agregar os casos tratados em função de um determinado critério;
- Um esquema de ponderação (pesos específicos) que permita valorizar os diversos produtos;
- Uma equação que possa ser aplicada à proporção de casos e aos respectivos pesos específicos.

Existem ainda diversas dimensões da produção hospitalar que podem servir de base para se estabelecerem os pesos específicos:

- A demora média por produtos dos hospitais (Schweiker, 1982);
- Os custos por produtos dos hospitais (Pettengill e Vertrees, 1980 e 1982);
- A concentração/especialização dos hospitais (Evans e Walker, 1972; Horn e Schumaker, 1979; Barer, 1982; McGuire e Williams, 1986 e Farley e Hogan, 1990);
- A gravidade dos casos tratados (Lichtig, 1986 e McGuire, 1991).

Como principais vantagens desta abordagem são referidas a sua simplicidade, visto que sintetiza num único indicador a actividade hospitalar e a minimização de alguns problemas técnicos e operacionais, como a multicolineariedade (Hornbrook, 1982).

No entanto, este mesmo autor identifica alguns problemas com a utilização de índices escalares (Hornbrook, 1982):

- Perda de informação e de sensibilidade derivada do nível de agregação existente;
- Possibilidade de existência de diferentes índices escalares sobre a actividade do hospital, em função da utilização de diferentes dimensões para se atribuírem pesos específicos (por exemplo, complexidade ou gravidade dos casos tratados);
- Possibilidade de não adequação para o financiamento dos hospitais, essencialmente se reflectirem as semelhanças e as diferenças entre os hospitais;
- Inadequação para a avaliação dos cuidados prestados, dado que a agregação existente no seu cálculo não permite distinguir entre cuidados apropriados e inapropriados.

2.2. Perfil de Admissões Hospitalares

As admissões hospitalares devem ser perspectivadas em função dos objectivos do hospital, os quais por sua vez podem apresentar objectivos contraditórios, como a optimização da capacidade instalada, a minimização da duração de internamento ou a satisfação dos clientes (Gemmel e Van Dierdonck, 1999).

Por outro lado, esta política de admissões hospitalares está intimamente associada com insuficiências do mercado da saúde, como a assimetria da informação existente entre os diversos agentes (Shen e Ellis, 2002), como a procura derivada e relação de agência (McGuire, 1989) e ainda com as especificidades das organizações de saúde, nomeadamente ao nível do seu modelo económico (Jacobs, 1974) e da existência de duas linhas de autoridade, com finalidades e funções de utilidade distintas (Costa e Reis, 1993).

Embora por razões diferentes o sistema de financiamento dos hospitais tem sido igualmente referido como um dos aspectos que poderá condicionar a política de admissões dos hospitais, nomeadamente quando se está na presença do pagamento prospectivo por caso (utilização dos DRGs, por exemplo), em que se espera que os hospitais tenham um incentivo para admitir casos menos graves dentro do mesmo produto (Hornbrook e Goldfarb, 1983; Gonnella, Hornbrook e Louis, 1984; Schumacher et al, 1987; Thomas e Ascraft, 1989 e Costa, 1990).

Todos estes aspectos têm conduzido à realização de diversos estudos sobre a selecção adversa nos serviços de saúde (Scitovsky, McNall e Benham, 1978; Ellis, 1998; Robinson e Gardener, 1995; Marquis e Buchanan, 1999; Shen e Ellis, 2002; Altman, Cutler e Zechauser, 2003; Meenan et al, 2003; Cao e McGuire, 2003 e Shapiro et al, 2003) nos quais se evidencia que os seguradores/hospitais tendem a escolher os doentes mais favoráveis.

A selecção adversa pode ser definida como o facto de alguns doentes, por exemplo, os mais idosos ou com doenças crónicas ou com resultados previsíveis mais desfavoráveis, serem rejeitados por pagadores ou por prestadores (Pereira, 1993).

No entanto, atendendo a que neste estudo não serão abordadas as questões mais específicas sobre a existência ou não de selecção adversa nos hospitais portugueses, passa-se de seguida a identificar e a caracterizar algumas dimensões qualificativas das admissões hospitalares.

Entre estas dimensões qualificativas do perfil das admissões hospitalares destacam-se as relacionadas com a concentração/especialização da produção hospitalar, com a complexidade dos casos tratados e com a gravidade dos casos tratados.

A concentração da produção num hospital pode ser definida pela proporção de produtos diferentes tratados em cada hospital (Hornbrook, 1982).

As principais razões para se estudar a concentração da produção são de dupla natureza: em primeiro lugar, avaliar se existe um comportamento distinto por parte dos hospitais, e na eventualidade de tal ocorrer, se estes comportamentos estão relacionados com algumas características dos hospitais; em segundo lugar, porque existe alguma evidência na literatura económica que a concentração/especialização da produção pode ter alguma relação directa com o desempenho dos hospitais, tanto em termos de eficiência, como da efectividade dos cuidados prestados (Horn e Schumaker, 1979; Watts e Klastorin, 1980; Goldfarb, Hornbrook e Higgins, 1983; Hornbrook e Monheit, 1985; McGuire e Williams, 1986; Farley e Hogan, 1990; Iezzoni et al, 1992b; Gowrisankaran e Town, 2003 e Shen, 2003).

A este propósito convém ainda distinguir a concentração da produção num determinado hospital da concentração da produção entre hospitais, visto que neste indicador se pretende medir a concentração de determinados produtos num número pequeno de hospitais (Schumaker et al, 1979).

Esta última medida de concentração de determinados produtos num número pequeno de hospitais baseia-se no princípio de que existe uma distribuição desigual dos produtos por hospital (Evans e Walker, 1972).

A sua operacionalização foi feita através do Índice da Teoria da Informação (Evans e Walker, 1972; Horn e Schumacher, 1979 e McGuire e Williams, 1986).

No estudo de Evans e Walker (1972) a medida de concentração é igualmente a medida da complexidade dos casos tratados, embora Tatchell (1983) refira que por vezes a complexidade dos casos é confundida com a raridade dos casos.

Por sua vez, Horn e Schumacher (1979) referem que a utilização do Índice da Teoria da Informação apresenta as seguintes vantagens: (1) é fácil de calcular, visto que a distribuição das doenças é conhecida; (2) é menos influenciada pelo hospital do que as medidas de complexidade, baseadas nos custos ou na demora média e (3) pode ser utilizado como variável independente em equações em que o custo médio por hospital é a variável dependente.

No entanto, McGuire e Williams (1986) na sua aplicação da Teoria da Informação às funções de custo dos hospitais escoceses encontraram resultados piores que os referidos nos estudos anteriores, argumentando que tal se poderá dever à própria medida, a qual poderá não estar a caracterizar em toda a extensão a produção hospitalar.

Por outro lado, Hornbrook (1982), embora refira que a Teoria da Informação possa constituir uma abordagem útil e válida, deve ser interpretada com algumas

cautelas, visto que a caracterização na desigualdade de distribuição de produtos entre hospitais, só por si não justifica se a mesma é desejável, indesejável ou irrelevante e ainda porque um indicador único pode conduzir a resultados enviesados ou imprecisos, visto que não identifica quais as áreas de especialização de cada hospital.

Atendendo a estes aspectos, parece mais relevante apurar a concentração da produção num hospital, que pode ser medida pela percentagem de produtos diferentes responsáveis pela totalidade da produção e por grande parte da produção (80% do total da produção total, por exemplo).

Estes aspectos embora não eliminem as críticas respeitantes ao carácter unidimensional do indicador e só por si não caracterizem as áreas de concentração ou de dispersão dos hospitais, permitem identificar eventuais comportamentos distintos dos hospitais.

A complexidade dos casos é um indicador que deve sintetizar a quantidade de recursos necessários para tratar cada doente (Luke, 1979). Este mesmo autor introduz ainda o conceito de intensidade na utilização de recursos, o qual permite medir este mesmo consumo de recursos, mas por dia de internamento (Luke, 1979).

Este conceito é um indicador que pretende caracterizar a diversidade da produção hospitalar com base em critérios de oferta, visto que será natural esperar que a complexidade de casos em determinado hospital esteja associada com o seu nível de desenvolvimento tecnológico e com a especialização dos recursos humanos (Hornbrook, 1982 e Becker e Steinwald, 1991).

Para se tornar operacional este conceito, para além do Índice da Teoria da Informação descrito anteriormente, pode recorrer-se igualmente a informação dos sistemas de classificação de doentes.

Entre estes podem destacar-se os Diagnosis Related Groups (DRGs) os quais pretendem definir produtos com consumo de recursos homogéneo, sendo o consumo de recursos medido pela duração de internamento (Fetter et al, 1980).

Por outro lado, com a utilização dos DRGs para finalidades de financiamento, foram simultaneamente definidos preços e pesos relativos por cada DRG, podendo estes últimos ser considerados como uma boa medida da complexidade dos casos tratados (McGuire, 1991).

A este propósito convém ainda referir uma outra noção praticamente singular de complexidade apresentada por Naessens e outros (1992), os quais definem a complexidade do caso como a extensão da gravidade não só ao diagnóstico principal, mas também à gravidade das comorbilidades existentes, tendo para tal utilizado o Disease Staging.

No entanto, atendendo às noções de complexidade geralmente aceites, bem como às definições de gravidade existentes, parece que aquela medida de complexidade é mais uma medida de gravidade, a qual pretende somente evidenciar a importância das comorbilidades.

Para a gravidade da doença ou severidade do estado do doente embora exista uma grande variedade de definições que passam desde o risco de morte, à dificuldade para o tratamento ou à instabilidade clínica (Iezzoni, 1997b) é pacificamente aceite que esta dimensão qualifica as características dos doentes e que um conceito aceitável é “Probabilidade de Morte ou de Falência de um Órgão” (Thomas, Ashcraft e Zimmerman, 1986).

Ao nível conceptual estão identificadas uma série de dimensões do risco do doente e consequentemente da gravidade do caso: Idade, Diagnóstico Principal, Comorbilidades, Estatuto Funcional, Atributos Culturais, Étnicos e Socioeconómicos e Atitudes e Preferências dos Consumidores (Blumberg, 1986; Gustafson et al 1986; Charlson et al, 1987; Zimmerman, 1989; Deyo, Cherkin e Ciol, 1992; Iezzoni et al, 1994a; Davis et al, 1995; Iezzoni, 1995; Hornbrook e Goodman, 1986 e Cleves, Sanchez e Drahein, 1997).

Para se tornar operacional este conceito existe igualmente uma enorme variedade de sistemas de classificação de doentes validados para os doentes internados, como o Acuity Index Method, os All Patient Refined DRGs (APRDRGs), o Computerized Severity of Illness, o Disease Staging, os International Refined DRGs (IRDRGs) e os MedisGroups (Thomas, Ashcraft e Zimmerman, 1986 e Iezzoni, 1997c). Para as Unidades de Cuidados Intensivos existem igualmente alguns sistemas de classificação validados, como o APACHE III (Zimmerman, 1989), o MPM II (Lemeshow et al, 1993) e o SAPS II (Le Gall, Lemeshow e Saulnier, 1993).

Face ao exposto, essencialmente que a complexidade constitui uma dimensão que privilegia as características da oferta para se qualificarem os produtos, enquanto que a gravidade é uma dimensão em que predominam as características dos doentes para se definirem estes mesmos produtos hospitalares, torna-se importante avaliar o perfil das admissões hospitalares em função destas duas perspectivas.

Mais uma vez se recorda que, neste momento, atentas as características do Serviço Nacional de Saúde (SNS) português onde avultam a universalidade e a definição de áreas de influência por hospital, mais relevante que identificar políticas de escolha de doentes por parte dos hospitais (selecção adversa), é a avaliação de eventuais comportamentos distintos entre estas organizações de saúde.

2.3. Avaliação do Desempenho Hospitalar

2.3.1. Importância do Tema

A avaliação do desempenho das organizações de saúde e mais concretamente dos hospitais é um assunto que tem merecido crescente interesse e importância, tanto em termos internacionais, como em Portugal.

Em termos internacionais podem citar-se, a título perfeitamente exemplificativo as experiências dos EUA – “Best Hospitals” (Hill, Winfrey e Rudolph, 1997; O’Muircheartaigh, Murphy e Moore, 2002), “One Hundred Top Hospitals” (HCIA, 1999; Griffith, Alexander e Jelinek, 2002) e “Risk-Adjusted Quality Outcomes Measures” (DesHarnais et al, 1997 e 2000), do Reino Unido (Amaratunga et al, 2002, Chang, Lin e Northcott, 2002; Department of Health, 2002; NHS, 2002 e Snelling, 2003) e da Austrália (Ibrahim et al, 1998; Degeling et al, 2000; NHPC, 2001 e ACHS, 2002).

Em Portugal, embora a situação seja mais incipiente são de destacar os trabalhos realizados pelo INA – “Avaliação dos hospitais Fernando Fonseca e Garcia da Orta”, 1999 e “Projecto de Avaliação de Unidades de Saúde”, 2001, e diversos estudos realizados pelo IGIF (Barros, 2001a; Barros 2001b e Barros, 2001c), bem como a experiência mais recente na avaliação dos hospitais SA. Inclusivamente ao nível académico, existem alguns estudos, dos quais se podem referir os realizados por Costa e Reis (1993), Dias Alves (1994), Dismuke e Sena (1998), Barros e Sena (1999), Cabral e Barriga (1999) e Carreira (1999).

A todos estes aspectos acrescem algumas experiências de avaliação da actividade dos hospitais realizadas no âmbito das Agências de Contratualização dos Serviços de Saúde (Contratualização com os Hospitais para 2003, 2002 e Projecto de Indicadores, 2002) e ainda da Direcção Geral de Saúde (Desempenho Comparado das Unidades de Saúde do SNS – Indicadores Mensais, 2002).

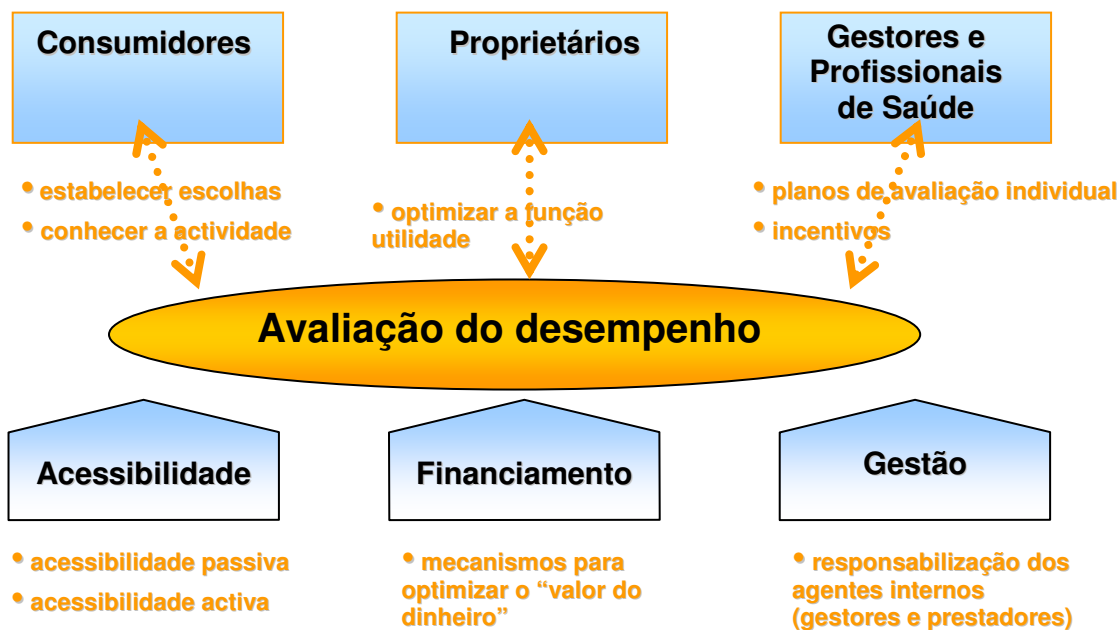
Ainda em relação a Portugal, a avaliação do desempenho das organizações de saúde parece ganhar ainda mais relevo após a publicação do “Novo Regime Jurídico da Gestão Hospitalar” (Lei nº 27/2002 de 8 de Novembro).

Neste diploma são de referir desde já os aspectos presentes no Art. 8º (Informação Pública) – “O Ministério da Saúde divulga, anualmente, um relatório com os resultados da avaliação dos hospitais que integram a rede de prestação de cuidados de saúde mediante um conjunto de indicadores que evidencie o seu desempenho e eficiência”, na alínea b) do Art. 5º (Princípios Específicos da Gestão Hospitalar) – “Garantia aos utentes da prestação de cuidados de saúde de qualidade com um controlo rigoroso dos recursos” e nas alíneas a), b) e c) do nº1 do Art. 10º (Princípios Específicos da Gestão Hospitalar do Sector Público

Administrativo) – a) “Garantia da eficiente utilização da capacidade instalada, designadamente pelo pleno aproveitamento dos equipamentos e infra-estruturas existentes e pela diversificação do regime de horário de trabalho, de modo a alcançar uma taxa óptima da utilização dos recursos disponíveis”, b) “Elaboração de planos anuais e plurianuais e celebração de contratos-programa com a Administração Regional de Saúde (ARS) respectiva, de acordo com o princípio contido na alínea d) do artigo 5º, nos quais sejam definidos os objectivos a atingir e acordados com a tutela, e os indicadores de actividade que permitam aferir o desempenho das respectivas unidades e equipamentos de gestão” e c) “Avaliação dos titulares dos órgãos de administração, dos directores dos departamentos e de serviços e dos restantes profissionais, de acordo com o mérito do seu desempenho, sendo este aferido pela eficiência demonstrada na gestão dos recursos e pela qualidade dos cuidados prestados aos utentes”.

Por outro lado, a dificuldade para se avaliar o desempenho das organizações de saúde tem sido evidenciado em diversos artigos (Anthony e Herzlinger, 1975; Evans, 1981 e Costa e Reis, 1993), sendo de referir os aspectos associados com a proliferação de perspectivas e de agentes, a existência de uma dupla linha de autoridade e a especificidade na garantia da qualidade dos cuidados prestados.

Figura 1
Necessidade de Avaliação das Organizações de Saúde (1)



No entanto, embora tendo patentes as dificuldades conceptuais e operacionais para a sua avaliação, torna-se cada vez mais necessário desenvolver modelos para a sua concretização. Esta afirmação é válida para qualquer dos agentes associados ao mercado da saúde (ver Figura 1).

Para os consumidores, visto que é cada vez mais imperiosa a disponibilização de informação científica e credível que lhes permita estabelecer escolhas e simultaneamente conhecer a actividade das organizações de saúde.

Para os proprietários, já que somente a explicitação de modelos válidos e “universais” lhes permitirá otimizar a sua função de utilidade.

Para os gestores e profissionais de saúde, porque a existência de um modelo de avaliação do desempenho das organizações de saúde, permite igualmente a sua extensão para o plano de avaliação individual e desta forma explicitar mecanismos de avaliação da competência e a possível atribuição de incentivos.

Para além destes aspectos sectoriais existe igualmente uma variedade de razões que tornam igualmente imperiosa a definição e implementação de modelos de avaliação do desempenho das organizações de saúde.

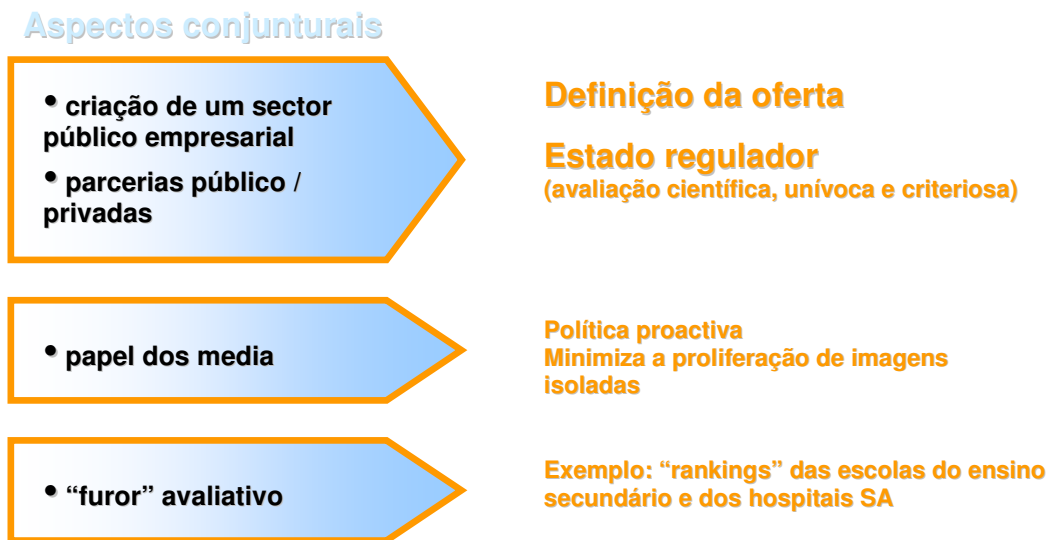
Em primeiro lugar, devem ser referidos os aspectos relacionados com a acessibilidade. De facto, este aspecto pode assumir diversas perspectivas, em conformidade com o enquadramento organizacional do sistema de saúde. Em Portugal, para além do imperativo constitucional que lhe está associado, este aspecto assume importância de dupla natureza, tanto associado à política de selecção/desnatação praticada por cada hospital (acessibilidade passiva), como no que se refere à disponibilização de informação que permita aos consumidores uma escolha fundamentada dos prestadores (acessibilidade activa).

Em segundo lugar, devem ser consideradas as questões associadas com o financiamento das organizações de saúde. A este propósito deve referir-se que independentemente da forma como será considerada a sua prospectividade – orçamento global, capitação ou pagamento pela produção – somente a definição prévia de mecanismos de avaliação do desempenho poderá de facto criar mecanismos para otimizar o “valor do dinheiro”.

Finalmente, em termos de gestão das organizações de saúde. Na realidade, somente após a definição, explicitação e implementação de um modelo de avaliação do desempenho das organizações de saúde, se poderão criar mecanismos que proporcionem uma maior responsabilização dos agentes internos (gestores e prestadores) e simultaneamente possibilitem a sua reprodutibilidade na actividade e gestão de cada organização de saúde.

Para além disso, existem em Portugal alguns aspectos conjunturais que tornam igualmente imperiosa a explicitação de modelos de avaliação do desempenho (ver Figura 2).

Figura 2
Necessidade de Avaliação das Organizações de Saúde (2)



Um primeiro grupo de razões está associado aos projectos de reforma existentes. De facto, tanto a criação de um sector público empresarial na saúde, como de parcerias público-privado para a construção de novas unidades, constituem aspectos críticos, tanto para a oferta – somente após a definição de um cenário de avaliação se podem estabelecer modelos previsionais para análises de rendibilidade dos capitais e dos recursos - como para o estado regulador – somente a criação destes mecanismos de avaliação lhe poderá permitir uma avaliação científica, unívoca e criteriosa da bondade das novas experiências de gestão.

Um segundo grupo de razões poderá ser encontrado na sociedade portuguesa. Uma perspectiva está associada à maior exposição que todos sofrem por parte da comunicação social. Neste sentido, pensa-se que a criação e divulgação de modelos de avaliação do desempenho permitirá uma política mais proactiva e poderá minimizar a proliferação de imagens isoladas, com grande predominância dos aspectos negativos e de visões parcelares, as quais actualmente constituem o alvo na avaliação do funcionamento das organizações de saúde. A outra perspectiva, mais global é encontrada na comparação com os

outros sectores de actividade. É praticamente generalizado o “furor” avaliativo e muitas vezes classificativo existente, sendo um bom exemplo disso a publicação dos “rankings” das escolas do ensino secundário. Inclusivamente no sector da saúde a recente publicitação de alguns elementos sobre o desempenho dos hospitais, nomeadamente os empresarializados, constitui igualmente um bom exemplo ilustrativo desta situação.

Atendendo a estes comportamentos e atitudes será natural considerar a eventualidade de a explicitação de modelos de avaliação do desempenho começar a ser um imperativo para o sector da saúde.

Tendo em vista os aspectos enunciados, qual deverá a ser o caminho a percorrer em Portugal?

Face ao exposto, pese embora o facto de as experiências e expectativas citadas assumirem características completamente diferentes, podem ser formuladas as seguintes questões:

- Quais são as dimensões e perspectivas de avaliação do desempenho dos hospitais?
- Quais são os destinatários da avaliação do desempenho dos hospitais?
- Que instrumentos e medidas devem ser utilizados para a avaliação do desempenho dos hospitais?

2.3.2. Algumas Experiências Internacionais

Na grande maioria das organizações o seu desempenho é avaliado pelos resultados (Costa e Reis, 1993). Nas organizações não lucrativas em geral e nas organizações de saúde este axioma tem sido posto em causa (Evans, 1981; McGuire, Henderson e Money, 1988).

Mas, será que esta situação deverá ser considerada estrutural (ponto de partida), ou antes pelo contrário, deverá ser perspectivada como uma consequência das dificuldades operacionais e conjunturais das organizações de saúde (ponto de chegada)?

A resposta a esta questão não é fácil, nem imediata.

Em Portugal, até ao presente momento, a abordagem dos resultados para a avaliação do desempenho das organizações de saúde não tem sido privilegiada (INA, 1999; INA, 2001; Barros, 2001a; Barros, 2001b; Barros, 2001c; DGS, 2002).

Por outro lado, a Lei nº 27/2002 de 8 de Novembro, embora acentue aspectos como a informação ao público e a avaliação da eficiência e da efectividade dos cuidados prestados é omissa em relação ao plano de avaliação.

O estado da arte da avaliação do desempenho das organizações de saúde parece seguir a trilogia definida por Donabedian (1985) – Estrutura, Processo e Resultados (ver Figura 3).

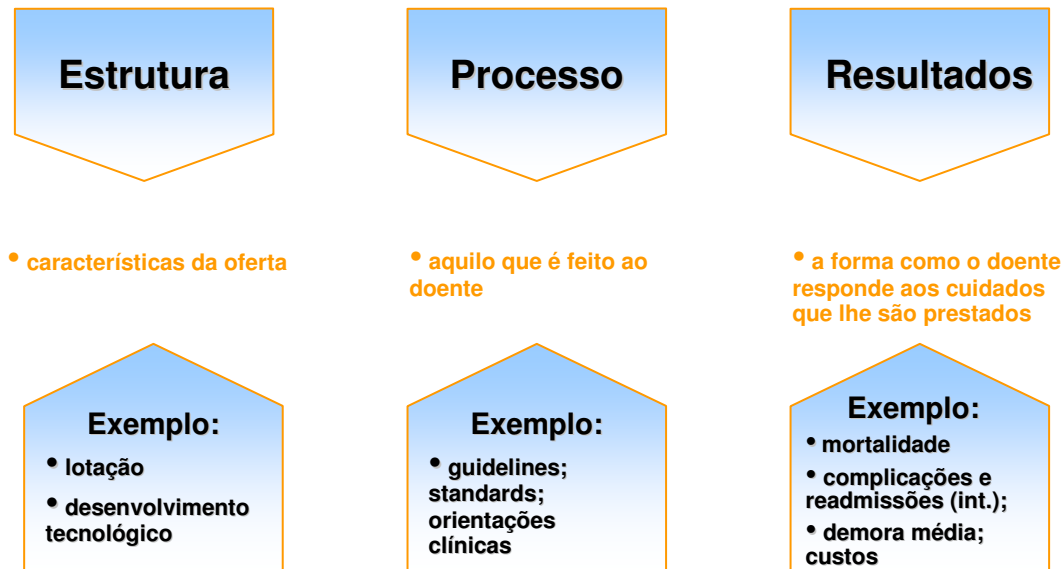
Em que estrutura pode ser definida pelas características da oferta (organizações de saúde), processo por aquilo que é feito ao doente e resultados pela forma como o doente responde aos cuidados que lhe são prestados (QMAS, 1997).

Na realidade embora esta trilogia tivesse originalmente sido concebida para aplicação na garantia da qualidade dos cuidados prestados, parece igualmente legítimo afirmar que este âmbito é aplicável a outras realidades da actividade hospitalar, designadamente na eficiência e também no desempenho global dos hospitais.

Nas experiências anteriormente referidas dos EUA, por exemplo, deve referir-se que enquanto no “Best Hospitals” (Hill, Winfrey e Rudolph, 1997; O’Muircheartaigh, Murphy e Moore, 2002), e no modelo preconizado por DesHarnais e colegas (1997 e 2000) se privilegia essencialmente a dimensão qualidade, com a particularidade de a primeira incluir indicadores de estrutura, do processo e dos resultados, enquanto que a segunda foca essencialmente os resultados.

Pelo contrário, o modelo desenvolvido pela HCIA (HCIA, 1999 e Griffith, Alexander e Jelinek, 2002), envolve as perspectivas da qualidade, eficiência e rentabilidade dos capitais, embora considere somente uma dimensão – os resultados.

Figura 3
Estado de Arte da Avaliação do Desempenho



No Reino Unido, existem basicamente duas grandes abordagens, a primeira baseada na definição de indicadores do Serviço Nacional de Saúde (Chang, Lin e Northcott, 2002), a segunda interessada na concretização de uma metodologia para estabelecer “hospital ratings” (NHS, 2002).

Qualquer das perspectivas referidas engloba indicadores de estrutura, de processo e de resultados.

Na Austrália, embora os modelos referidos sejam completamente diferentes, oscilando entre indicadores de qualidade (Ibrahim et al, 1998 e AHCS, 2002), e outros mais gerais, com indicadores sobre financiamento, acesso, eficiência e qualidade (Degeling et al, 2000 e NHPC, 2001) a perspectiva privilegiada nos dois tipos de modelos é o processo.

Em seguida irão ser genericamente descritas algumas metodologias utilizadas nos EUA, atendendo essencialmente ao grau de aplicação de sistemas de

avaliação do desempenho dos hospitais neste país e no Reino Unido devido à maior semelhança do seu modelo organizacional com a realidade portuguesa.

A metodologia preconizada pelo “Best Hospitals”, começou a ser aplicada nos EUA em 1991, tendo sido desenvolvida pelo National Opinion Research Center (NORC) da Universidade de Chicago, tendo como principal finalidade a classificação dos hospitais norte-americanos. Para tal são utilizados dados que incluem alguns aspectos da estrutura, do processo e dos resultados.

Com a combinação destes indicadores pretende-se construir um Índice de Qualidade Hospitalar, seguindo basicamente a seguinte metodologia (O’Muircheartaigh, Murphy e Moore, 2002):

- Em primeiro lugar é definido o conjunto de hospitais que será incluído na análise. A metodologia é iterativa, constituindo o primeiro passo na definição de aspectos qualitativos, os quais incluem essencialmente elementos relacionados com a função ensino e com o grau de desenvolvimento tecnológico dos hospitais. O segundo passo, constitui na aplicação de critérios quantitativos – mínimo de doentes tratados, total e casos cirúrgicos, por especialidade – para o sub-conjunto de hospitais previamente escolhido;
- Em seguida é definida a metodologia para se operacionalizar a medição de cada uma das três dimensões da qualidade dos cuidados prestados. Para a estrutura são utilizadas 11 medidas, as quais compreendem aspectos como o desenvolvimento tecnológico, a qualificação dos recursos humanos, o volume, a diversificação e a qualificação da produção e o planeamento da continuidade de cuidados. A estas variáveis são atribuídos pesos relativos específicos por especialidade. Para o processo é utilizada uma medida indirecta – a reputação dos hospitais – que é calculada através da elaboração de um questionário destinado a um painel nacional de peritos médicos por especialidade. Para os resultados é utilizada a comparação entre mortalidade observada e esperada, sendo utilizados os All Patient Refined – Diagnosis Related Groups (APRDRGs), como medida de ajustamento do risco;
- Finalmente, o Índice de Qualidade Hospitalar para cada organização é calculado com a atribuição de pesos específicos idênticos para as três dimensões de qualidade.

Após o apuramento do Índice de Qualidade Hospitalar existem dois tipos de classificações. Uma primeira, genérica, na qual os hospitais são avaliados e ordenados pelo seu desempenho global. A segunda é efectuada por especialidade, sendo igualmente identificados os melhores hospitais.

O sistema desenvolvido por DesHarnais e colegas (1997 e 2000) é genericamente conhecido por “Risk Adjusted Quality Outcome Measures” e pretende construir um índice para medir a qualidade dos cuidados prestados.

Para tal foram consideradas três dimensões da qualidade: mortalidade, complicações e readmissões.

Para a mortalidade foi construído um indicador “Risk-Adjusted Mortality Index” (RAMI). Para o cálculo utilizou-se a regressão logística, na qual as variáveis explicativas são a idade, sexo, raça, o “cluster” do DRG, presença ou ausência de comorbilidades, presença de qualquer diagnóstico secundário de cancro, com excepção do cancro da pele e o número total das comorbilidades.

Para as complicações definiu-se o “Risk-Adjusted Complications Index” (RACI), utilizando-se mais uma vez a regressão logística, onde as variáveis explicativas são a idade, sexo, o “cluster” do DRG, presença ou ausência de comorbilidades, presença de qualquer diagnóstico secundário de cancro, com excepção do cancro da pele e o número total das comorbilidades.

Para as readmissões construiu-se o “Risk-Adjusted Readmissions Index” (RARI). Na regressão logística utilizaram-se como variáveis explicativas a idade, sexo, o “cluster” do DRG, presença ou ausência de comorbilidades e complicações, presença de qualquer diagnóstico secundário de cancro, com excepção do cancro da pele e o número total das comorbilidades.

Os três indicadores são utilizados separadamente, pelo que não foi concebido um indicador global de qualidade de cuidados prestados.

O modelo desenvolvido pela HCIA “One Hundred Top Hospitals” (HCIA, 1999; Griffith, Alexander e Jelinek, 2002) considera quatro dimensões de avaliação: (1) Financeiro, (2) Processo Interno, (3) Satisfação do Cliente e (4) Crescimento e Aprendizagem.

No Plano Financeiro pretende-se avaliar o respectivo desempenho dos hospitais, bem como a gestão de recursos. São utilizadas duas medidas:

- “Cash Flow Margin” – pretende descrever o grau de “lucro” das organizações. É calculado pelo somatório dos lucros, amortizações e despesas com juros dividido pelo total líquido das receitas;
- “Asset Turnover” – total dos activos dividido pelo total líquido das receitas.

No Processo Interno pretende-se avaliar o custo, a qualidade, a eficiência e outras características da produção hospitalar. São utilizadas três medidas:

- “Mortality Index” – A mortalidade observada é dividida pela mortalidade esperada, sendo utilizados valores de 2 anos. Este valor é calculado por diagnóstico principal através de regressão logística, na qual as variáveis explicativas são a idade, sexo, procedimentos efectuados, comorbilidades e a localização, dimensão, estatuto de ensino e estatuto urbano/rural dos hospitais;
- “Complications Index” – Avalia-se a incidência de 43 indicadores negativos sobre infecções pós-operatórias, hemorragias pós-operatórias ou infecções nosocomiais nas complicações. Este índice é ajustado pelo risco, com uma metodologia idêntica à usada para a mortalidade;
- “Cost per Case” – Custos por caso, ajustados para os cuidados prestados em ambulatório e em situações não agudas pelo “casemix” e pelos níveis salariais.

Na Satisfação do Cliente pretendem-se avaliar aspectos como a quota de mercado e a situação competitiva dos hospitais. São utilizadas duas medidas:

- “Occupancy” – É calculada pela divisão do total de dias de internamento pela capacidade total (Taxa de Ocupação). É uma medida indirecta da quota de mercado;
- “Changes in Occupancy” – É calculado pela diferença entre a Taxa de Ocupação do presente ano e a do ano anterior dividida pela Taxa de Ocupação do presente ano. É uma medida indirecta da competitividade do hospital.

No Crescimento e Aprendizagem pretende-se avaliar a capacidade de cada organização para responder às alterações tecnológicas, às preferências dos consumidores e ao ambiente competitivo. São utilizadas duas medidas:

- “Lenght of Stay” – É calculado pela divisão do total de dias de internamento dividido pelo total de doentes tratados (Demora Média). Mais uma vez, os resultados dos hospitais no que respeita à demora média são ajustados pela gravidade. Pretende dar indicações sobre o nível de eficiência dos hospitais e sobre a redução de riscos nos doentes;
- “Outpatient Activity” – É calculado pela divisão do total de receitas líquidas em ambulatório pelo total de receitas líquidas. É uma medida indirecta das actividades preventivas dos hospitais.

Este tipo de análise é efectuado em cinco grupos diferentes de hospitais:

- Hospitais Pequenos – Entre 25 e 99 camas
- Hospitais Médios – Entre 100 e 249 camas
- Hospitais Grandes – Mais de 250 camas

- Hospitais Universitários – Hospitais com 250 ou mais camas e com ratios de 5 internos por médico ou ratios de interno por médico e por cama entre 0.01 e 0.24
- Hospitais Universitários Grandes – Hospitais com 400 ou mais camas e um ratio interno por médico e por cama superior a 0.25.

Todas as medidas são utilizadas de forma isolada para se compararem hospitais dentro dos grupos definidos.

No Reino Unido, uma primeira abordagem consiste na definição de indicadores do Serviço Nacional de Saúde (Chang, Lin e Northcott, 2002 e NHS, 2002), conforme consta do Quadro I. Esta metodologia compreende duas grandes áreas: o desempenho do Serviço Nacional de Saúde e o desempenho dos hospitais “trusts” do Serviço Nacional de Saúde.

Quadro I
 Reino Unido: Lista de Indicadores do Desempenho Hospitalar

I (a) Efectividade Clínica e Resultados	II Eficiência	III Experiência dos Doentes	IV Dimensão e Capacidade
% de doentes com mais de 50 anos, com alta para casa dentro de 56 dias, após a admissão por doença cerebrovascular	Ratio entre internamentos observados e esperados para um conjunto de 25 procedimentos, ajustados pelo “casemix”	% de doentes com espera para internamento inferiores a 6 meses	% de internos em função das normas do internato
% de doentes com mais de 65 anos, com alta para casa dentro de 28 dias, após a admissão por fractura da anca	Ratio entre a demora média observada e esperada, ajustada pelo “casemix”	% de doentes com consulta hospitalar dentro de 13 semanas, após a referência do médico de família	Vagas no preenchimento dos quadros em relação ao total de efectivos previsto - pessoal técnico de saúde
Readmissões dentro de 28 dias em relação ao total de altas vivos – para a fractura da anca	Comparação dos custos operacionais de cada hospital com os valores médios	% de doentes com consulta de especialidade dentro de 2 semanas após a referência do médico de família – cancro da mama	Vagas no preenchimento dos quadros em relação ao total de efectivos previsto – enfermeiros e parteiros
Readmissões dentro de 28 dias em relação ao total de altas vivos – para a doença cerebrovascular	% de 1 ^{as} consultas marcadas e não realizadas	% de doentes admitidos pela urgência sem cama após 4 horas da decisão de internamento	Vagas no preenchimento dos quadros em relação ao total de efectivos previsto – outro pessoal
Óbitos após 30 dias de admissão para situações não electivas, por 100,000 doentes		% de reclamações existentes para situações com resolução local dentro de 4 semanas	Taxa de absentismo
Óbitos após 30 dias de admissão para “bypass” coronário, por 100,000 doentes		% de intervenções cirúrgicas programadas canceladas	Nível de cumprimento do sistema de gestão do risco
Óbitos após 30 dias de admissão para fractura da anca (+65 anos), por 100,000 doentes		% de doentes não admitidos após um mês do cancelamento da intervenção cirúrgica	Medida de classificação da qualidade da informação do hospital – Resumo de Alta
Óbitos após 30 dias de admissão para a doença cerebrovascular, por 100,000 doentes		Pontuação global da higiene e limpeza do hospital	

Fonte: Chang, Lin e Northcott, 2002

(a) – todos os indicadores são padronizados por idade e sexo

Em relação ao Serviço Nacional de Saúde são consideradas seis dimensões:

- Melhoria da Situação de Saúde;
- Acessibilidade;
- Apropriação e Efectividade dos Cuidados Prestados;
- Eficiência;
- Experiência dos Doentes;
- Resultados de Saúde do SNS.

No que se refere aos hospitais são somente consideradas quatro dimensões:

- Efectividade Clínica e Resultados (I);
- Eficiência (II);
- Experiência dos Doentes (III);
- Dimensão e Capacidade (IV).

Para cada uma destas dimensões são considerados os indicadores referidos no Quadro seguinte. Todos os indicadores são apresentados isoladamente, não existindo qualquer sistema de classificação global.

Ainda no Reino Unido existe uma segunda abordagem interessada na concretização de uma metodologia para estabelecer “hospital ratings” (NHS, 2002 e Snelling, 2003).

Este sistema inclui três fases:

- Identificação de indicadores chave (“Key Targets”);
- Identificação de outras medidas relevantes;
- Revisão da actividade e classificação dos hospitais.

Os Indicadores Chave correspondem às medidas mais importantes no que se refere à atribuição da pontuação (“star rating”) e compreendem os aspectos considerados mais relevantes na prestação de cuidados de saúde. Por exemplo um “key target” é: “Nenhum doente pode esperar mais de 18 meses para receber tratamento hospitalar”. Considera-se que este é atingido quando nenhum doente está nesta condição, parcialmente atingido quando 1 ou 2 doentes esperam mais de 18 meses, ou falhado quando 3 ou mais doentes têm um tempo de espera superior.

As Outras Medidas Relevantes estão agrupadas em três categorias:

- Clínica (risco de negligência clínica, óbitos no período de 30 dias para cirurgias não programadas, óbitos no período de 30 dias para “bypass” cardíaco, por exemplo);

- Doente (total da lista de espera para internamento, consultas externas efectuadas até 13 semanas após referenciação, doentes não operados dentro de 1 mês após o cancelamento de uma intervenção cirúrgica, por exemplo);
- Capacidade (taxa de absentismo, cumprimento das normas e de procedimentos sobre confidencialidade dos dados e informação disponibilizada, satisfação profissional, por exemplo).

Para qualquer uma das medidas o valor do hospital é comparado com a média nacional, sendo atribuída uma classificação ordinal entre 1 (pior) e 5 (melhor).

A revisão da actividade e classificação dos hospitais deve ser feita anualmente e com a seguinte lógica:

- ☆☆☆ (3 estrelas) – Para obter 3 estrelas o hospital deve atingir todos os Indicadores Chave, ou falhar apenas 1 numa quantidade mínima. Deve ter uma situação acima da média na grande maioria dos Outros Indicadores.
- ☆☆ (2 estrelas) – Para obter 2 estrelas o hospital deve apresentar bom desempenho na maioria dos Indicadores Chave e Outros Indicadores. Ou se atingir todos os Indicadores Chave tiver mau desempenho nos Outros Indicadores ou ainda, se tiver um mau desempenho em poucos Indicadores Chave e muito bom desempenho nos Outros Indicadores.
- ☆ (1 estrela) – O hospital é classificado com 1 estrela se falhar vários Indicadores Chave, ou mesmo que apresente um bom desempenho nestes indicadores apresentar uma má situação (muitos indicadores abaixo da média) nos Outros Indicadores.
- ○ (sem estrelas) – Ao hospital é atribuída uma classificação sem estrelas se apresentar um mau desempenho na globalidade dos Indicadores Chave, mesmo que tenha boa classificação nos Outros Indicadores.

A análise dos sistemas de avaliação do desempenho anteriormente apresentados permite situar o respectivo “estado da arte” em dois planos: a dimensão da sua aplicação e a perspectiva de avaliação.

No que se refere à dimensão da aplicação de sistemas de avaliação do desempenho dos hospitais retoma-se a discussão anteriormente efectuada – estrutura, processo e resultados.

Independentemente da necessidade de se considerarem todas as dimensões para avaliar a actividade dos hospitais (Brook et al, 1987; Murphy e Cluff, 1990; Kanouse, Kallich e Kahan, 1995; Almeida e Carlsson, 1996; Iezzoni, 1997b e Weingart et al, 2001), a consideração da dimensão resultados parece assumir

cada vez mais importância, tanto pela aproximação que permite aos modelos utilizados na grande maioria das organizações do mercado em geral, como pela proliferação e utilização cada vez mais intensa nas organizações de saúde, em termos internacionais.

Em Portugal, este aspecto deve assumir cada vez maior relevo, tanto no que se refere à cultura organizacional emergente, da qual a avaliação das escolas e dos hospitais SA constituem um bom exemplo, como essencialmente atendendo às reformas a decorrer no sector da saúde, nas quais as alterações ao modelo de gestão hospitalar surgem como aspectos de partida e ainda nos aspectos decorrentes das maiores exigências ao nível da responsabilização e da “accountability” dos hospitais, pelo que se pode concluir que a avaliação dos resultados constitui um novo desafio à actividade reguladora do Estado.

Neste sentido, mesmo tendo em conta a necessidade de implementação de mecanismos e de metodologias que avaliem a estrutura e o processo das organizações de saúde, defende-se que a avaliação dos seus resultados constitui o elemento mais decisivo, assumindo-se que aos hospitais que comparativamente apresentem melhores valores, têm, no mínimo, de apresentar bons desempenhos nas dimensões que estão a montante.

Em relação à perspectiva de avaliação podem ser consideradas três vertentes – a efectividade dos cuidados prestados, a eficiência das organizações de saúde e o seu desempenho financeiro.

Mais uma vez, embora de forma e intensidade diferentes estes aspectos são considerados nos sistemas de avaliação do desempenho das organizações de saúde anteriormente descritos.

A este propósito a decisão sobre a perspectiva a considerar é mais complexa, tanto atendendo a aspectos estruturais, como conjunturais. De facto, em termos conceptuais, somente se pode considerar que um hospital apresenta um bom desempenho se simultaneamente cumprir aquelas três perspectivas, ao que acresce a situação actualmente existente em Portugal, na qual são frequentemente indicados problemas nos três níveis.

No entanto, pese embora o facto de o desempenho financeiro ser eventualmente considerado relevante, neste estudo somente serão analisadas as questões relacionadas com a efectividade e com a eficiência.

Esta decisão permite, por um lado dar maior transparência à actividade dos hospitais, desde que a informação seja disseminada pela sociedade e inclusivamente introduzir novos elementos que permitam uma decisão mais criteriosa dos consumidores sobre a organização de saúde que pretendem escolher.

2.3.3. Ajustamento pelo Risco – Conceptualização e Operacionalização

2.3.3.1. Razão de Ser

Independentemente da perspectiva de avaliação dos resultados – efectividade, eficiência ou desempenho financeiro – o ajustamento pelo risco é necessário para se avaliar a actividade das organizações de saúde, visto que é necessário medir as características dos doentes que podem influenciar os resultados de saúde.

Em termos genéricos a principal finalidade do ajustamento pelo risco é a de controlar os factores que os doentes apresentam ao contactar uma determinada organização de saúde que podem afectar a sua probabilidade de obterem um bom ou um mau resultado (lezzoni, 1996a). Assim, o “rationale” para o ajustamento pelo risco consiste na eliminação dos factores que podem conduzir a diferentes resultados de saúde, essencialmente aqueles que estão associados aos doentes (lezzoni, 1997b e 1997c).

Contudo, algumas questões permanecem ainda em grande debate (lezzoni, 1996b; lezzoni, 1997c; Thomas e Hofer, 1998 e Thomas e Hofer, 1999):

- O que é significa, exactamente, risco?
- Como é que o risco deve, exactamente, ser medido?

A primeira questão que se deve colocar respeita ao risco de quê? Para esta interrogação pode existir uma grande variedade de respostas, variando desde o consumo de recursos, para a probabilidade de morte iminente, para a existência de complicações ou ainda para o desempenho funcional dos doentes e para a satisfação do cliente.

Este aspecto conduz a duas situações. A primeira é que não existe uma única abordagem de ajustamento pelo risco que apresente o mesmo nível de significância para todos os diferentes resultados de saúde (lezzoni, 1996b).

A segunda é que existe uma grande variedade de abordagens que pretendem estabelecer medidas de ajustamento pelo risco (Costa, 1991; lezzoni, 1997c).

Por outro lado, o ajustamento pelo risco tem sido frequentemente associado à questão da severidade do estado do doente.

Este assunto é ainda mais importante quando se tem presente que não existe uma definição única sobre severidade/gravidade.

Na realidade, Gertman e Lowenstein (1984) sintetizam esta dificuldade na definição de severidade, avançando que severidade é aquilo que os sociólogos designam como uma expressão da sabedoria popular, como satisfação ou

felicidade, operacionalmente indefinível, mas facilmente aceite por todos os intervenientes.

Ou ainda, como referem Thomas e Longo (1989) que, ao executarem uma extensa revisão de literatura sobre a severidade encontraram definições tão distintas como probabilidade de morte, diminuição das funções, “carga da doença”, integridade fisiológica, ou dificuldade de repor o nível de saúde dos doentes.

Tudo isto tem levado a uma grande proliferação de sistemas de classificação de doentes que pretendem medir a severidade, cada qual com a sua definição (Costa, 1991; lezzoni, 1997b e lezzoni, 1997c). Ou ainda a que existam diversas perspectivas conforme os agentes, variando por especialidades médicas e entre médicos e gestores, estando os primeiros mais interessados nos resultados de saúde e os segundos com a necessidade de recursos e com a intensidade do consumo de recursos (lezzoni, 1995).

Assim, a primeira questão que se coloca quando se pretende definir o sistema de ajustamento pelo risco reside no seu âmbito de aplicação:

A definição de produtos deve incidir essencialmente sobre a capacidade de um sistema definir produtos homogéneos em relação ao consumo de recursos, em relação à significância clínica ou ainda em relação a ambos?

A resposta a esta questão parece, numa primeira abordagem, óbvia e simples: o sistema de classificação de doentes deve estabelecer categorias que simultaneamente proporcionem a sua homogeneidade clínica e económica.

Neste sentido, a criação de um sistema misto possibilita, não só a sua múltipla utilização, conforme as finalidades a que se destina, mas também uma potencial minimização dos conflitos entre as duas linhas de autoridade presentes no hospital, visto que o plano de acção e de análise se centra no mesmo instrumento.

Contudo, a operacionalização deste princípio tem suscitado diversos problemas e contradições.

Em primeiro lugar, porque conforme referem Hornbrook (1982) e Aronow (1988), entre outros, a definição prévia da finalidade principal a que um sistema se destina constitui quase exclusivamente a única metodologia objectiva para se avaliar a sua efectividade.

Em segundo lugar, atendendo a aspectos de uma natureza mais prática, essencialmente os que se referem às necessidades imediatas de utilização do sistema, para financiamento ou para garantia de qualidade, por exemplo, e ainda aos próprios custos de implementação e de exploração do sistema, têm

originado uma proliferação de sistemas de classificação dos doentes para se medir exactamente a mesma realidade.

Tudo isto tem conduzido a uma primazia do princípio da definição da finalidade, remetendo-se para um momento posterior a investigação sobre a adequação do sistema de classificação de doentes a outras finalidades.

Assim, não é de estranhar que a grande discussão que ainda existe no panorama internacional se situe sobre a predominância de uma perspectiva em relação à outra, com sistemas que privilegiam a utilização de recursos, como por exemplo os Diagnosis Related Groups (DRGs) ou o Acuity Index Method (AIM), outros a severidade como o Acute Physiologic and Chronic Health Evaluation (APACHE), o System of Acute Physiologic Score (SAPS), o Mortality Probability Model (MPM), o Disease Staging ou os MedisGroups, por exemplo, enquanto que outros afirmam que consubstanciam os dois princípios, como os Patient Management Categories (PMCs), por exemplo (Iezzoni, 1997b).

A definição e proliferação de sistemas de classificação de doentes para a medição da severidade do doente têm suscitado diversas questões, dentro das quais avultam os aspectos relacionados com as dimensões do risco, com a existência de associação entre os resultados de saúde (mortalidade, complicações e readmissões) e a qualidade dos cuidados prestados ou com a eficiência, a escolha de modelos de ajustamento pelo risco e as propriedades estatísticas dos modelos de ajustamento pelo risco.

Atendendo à sua importância conceptual e operacional são estes os aspectos que se passarão a analisar com maior profundidade.

2.3.3.2. Dimensões do Risco

Actualmente, é quase pacificamente aceite que as dimensões de risco são as seguintes (Knaus et al, 1985 e 1991; Blumberg, 1986; Horn, 1988; Murphy e Cluff, 1990; Le Gall, Lemeshow e Saulnier, 1993; Lemeshow et al, 1993 e Iezzoni, 1997d):

- Idade;
- Sexo;
- Estado fisiológico do doente;
- Diagnóstico principal;
- Gravidade do diagnóstico principal;
- Dimensão e gravidade das comorbilidades;
- Situação/estado (“status”) funcional;
- Situação psicológica e cognitiva dos doentes;
- Atributos culturais, éticos e socioeconómicos;
- Atitudes e preferências dos consumidores.

A conjugação destes factores permite igualmente afirmar que um determinado sistema de classificação de doentes tem significância clínica.

Definindo-se significância clínica (Wood, Ament e Kobrinsky, 1981) pela extensão e compreensão do conhecimento da situação de saúde de cada doente proporcionada pelo sistema de classificação, sendo estas adaptáveis às expectativas dos médicos, permitindo simultaneamente que, sem outros elementos adicionais, a troca de informações sobre essas mesmas expectativas.

Idade

A idade é uma característica intrínseca dos doentes que não depende da efectividade dos cuidados prestados. Por outro lado, é igualmente aceite que para a grande maioria das situações clínicas, os idosos apresentam maiores probabilidades de desenvolverem situações adversas que a população jovem (Iezzoni, 1997d).

Neste sentido, tem sido defendida a inclusão da idade como factor de risco, independentemente das outras características consideradas. Esta posição é bem patente na grande maioria dos sistemas desenvolvidos até ao momento (APACHE, MPM, ou SAPS, por exemplo), tendo inclusivamente sido evidenciada a sua ausência como justificativa de desempenhos inferiores de alguns sistemas de medição da severidade, como por exemplo é patente no estudo de Blumberg (1991) em que, utilizando os MedisGroups ficou provado que a consideração da idade aumentava o poder explicativo deste sistema.

O aumento da idade, não só aumenta o risco de morte, mas também a probabilidade de se desenvolverem complicações (Goldman et al, 1977; Weintraub et al, 1991 e Forman et al, 1992).

Poderão ainda ocorrer alguns problemas na consideração da efectividade média dos hospitais para a categoria dos muito idosos. Na realidade, embora subsista alguma polémica, alguns estudos evidenciam que os idosos são tratados com menos intensidade para as mesmas situações comparativamente com populações mais jovens (Greenfield et al, 1987 e Bennet et al, 1991).

Outro aspecto eventualmente a considerar no factor idade é a distinção entre recém-nascidos, crianças, adolescentes e adultos. De facto, para além da consideração de um diferente conjunto de variáveis explicativas para o risco de morte, podem ainda existir diferentes níveis de “normalidade” e de “anormalidade” para as mesmas variáveis em função da idade (Pollack, Ruttiman e Getson, 1988 e Pollack, Patel e Ruttiman, 1996).

Para obviar a esta situação alguns estudos utilizam somente as populações adultas (com 18 ou mais anos) para se apurar o risco de morte, como por exemplo se pode constatar nos artigos de Iezzoni e outros (1995a) e de Thomas, Ashcraft e Zimmerman (1986).

Sexo

As diferenças entre os sexos podem igualmente constituir um factor suplementar para se avaliar o risco de morte, tanto pelas diferenças anatómicas, fisiológicas e hormonais existentes, como pelo facto de existirem algumas situações em que a resposta à terapêutica é diferente entre os dois sexos (Iezzoni, 1997d).

Ainda neste sentido, um estudo realizado por Gordon e Rosenthal (1999), para além de identificarem diversos estudos com diferentes riscos entre o sexo masculino e o sexo feminino, acabam por concluir pela existência de um risco superior de morte para os homens em quatro doenças com tratamento médico (doença cerebrovascular, doenças respiratórias das vias superiores, hemorragia gastrointestinal e pneumonia), embora para o enfarte agudo do miocárdio e para a insuficiência cardíaca os riscos de morte sejam praticamente idênticos entre os dois sexos.

Contudo, a grande maioria dos sistemas desenvolvido até ao momento não inclui o sexo como factor de risco da morte, visto que o seu poder explicativo não tem sido significativo (Francks, Gold e McClancy, 1996 e Iezzoni, 1997d).

Estado fisiológico do doente

A avaliação da estabilidade clínica reflecte o funcionamento fisiológico em termos de equilíbrio do estado geral de saúde (Iezzoni, 1997d).

Esta avaliação do estado fisiológico é considerada de importância crucial na previsão da mortalidade a curto prazo, visto que estas variáveis representam as funções orgânicas que mantêm os doentes vivos (Iezzoni, 1997d).

É devido a esta natureza global que frequentemente se considera que o resultado de cada variável é independente do diagnóstico/doença, sendo consequentemente um bom indicador do risco de cada doente (Knaus, Wagner e Draper, 1984; Knaus et al, 1985 e 1993; Le Gall, Lemeshow e Saulnier, 1993 e Lemeshow et al, 1993).

Estes mesmos autores referem ainda que com poucas variáveis, as quais por sua vez constituem rotinas dos hospitais, se conseguem alcançar sistemas de classificação de doentes com significância clínica e igualmente com bons níveis de previsão do risco.

Apesar destas vantagens é igualmente referido que estas variáveis fisiológicas podem não constituir o único vector de previsão da mortalidade hospitalar. É essencialmente atendendo a este facto que sistemas como o APACHE ou MPM, por exemplo, incluem outro tipo de variáveis para complementar o modelo para

previsão da mortalidade, como é o caso das comorbilidades presentes no momento de admissão ou inclusivamente o tipo de contacto com o hospital (Knaus, Wagner e Draper, 1984; Knaus et al, 1985 e 1993; Le Gall, Lemeshow e Saulnier, 1993 e Lemeshow et al, 1993).

Por outro lado, pese embora o facto de a medição do estado fisiológico do doente poder ser global, existem algumas correntes que defendem a inclusão de escalões de variação por variáveis distintas em função do diagnóstico principal (Horn, 1986, 1988 e 1997) ou ainda a consideração de pesos específicos por diagnóstico (Daley et al, 1988).

Finalmente, deverá ser tomada em consideração a metodologia para escolha das variáveis representativas do estado fisiológico do doente.

Basicamente três alternativas se colocam (Gustafson et al, 1986 e Iezzoni, 1997b):

- Utilização de painéis de peritos;
- Utilização de painéis de peritos, com posterior investigação/modelização empírica;
- Investigação/modelização empírica.

Na prática as três alternativas têm sido utilizadas. Sistemas como o Disease Staging, PMCs, MedisGroups e APACHE I e II, por exemplo, recorreram a painéis de peritos para escolher e atribuir “scores” às variáveis. Sistemas como o APACHE III utilizaram painéis de peritos para escolha das variáveis e modelização empírica para atribuição dos “scores” por variável. Sistemas como o MPM utilizaram a modelização empírica para escolha e atribuição de “scores” por variável, tendo, num segundo momento, recorrido a painéis de peritos para discutir a pertinência de todas as variáveis incluídas e/ou a necessidade de incluir outras variáveis não incluídas no modelo.

O principal aspecto a salientar tem residido na transição histórica destas formulações. De facto, os primeiros sistemas a serem desenvolvidos foram baseados quase exclusivamente nas opiniões de painéis de peritos, existindo inclusivamente quem defendesse (Gustafson et al, 1986) que este tipo de metodologia era a única possível para a definição de sistemas de medição da severidade.

Mais recentemente o recurso à modelização empírica tem vindo a ganhar relevo e apoiantes, pelo que parece a lógica mais válida para o desenvolvimento de um modelo de previsão da mortalidade hospitalar.

Diagnóstico principal

Segundo Hornbrook (1982) o diagnóstico é uma hipótese em relação à natureza da doença, aplicando-se de seguida esta hipótese de diagnóstico no tratamento e na recuperação do doente. Neste sentido, os diagnósticos condicionam a tecnologia de tratamento e consequentemente o tipo e intensidade de cuidados necessários para tratar cada doença.

Contudo, a identificação do diagnóstico poderá somente constituir o primeiro passo na abordagem do doente, visto que o estabelecimento do prognóstico é igualmente importante na determinação do processo de tratamento de cada doente (Gonnella, Hornbrook e Louis, 1984).

Atendendo a que, o prognóstico de cada situação pode ser perspectivado a curto, a médio e longo prazo e a que, frequentemente, este mesmo prognóstico em função do tempo poderá ser diferente e concomitantemente operacionalizado de outra forma, interessa delimitar o âmbito de aplicação deste conceito (Hughes et al, 1996).

Como o principal objectivo desta metodologia consiste na identificação dos produtos do hospital para ajustamento pelo risco, nesta fase delimitada ao internamento, interessa muito mais considerar o prognóstico a curto prazo. Assim, a manutenção da vida tem constituído a maior preocupação nesta situação, pelo que o prognóstico é frequentemente considerado em termos de morte (Blumberg, 1986).

No entanto, a morte nem sempre constitui o episódio terminal para todos os diagnósticos, pelo que dificuldades adicionais emergem quando se considera o diagnóstico como uma das dimensões do risco (Hill, Winfrey e Rudolph, 1997).

Por outro lado, na comparação de risco entre diferentes diagnósticos, têm igualmente subsistido diversos problemas, tanto quando se comparam diagnósticos em que a morte é um “outcome” frequente, como quando a comparação é exercida entre diagnósticos com probabilidades de morte distintas.

De facto, a comparação da severidade entre doenças tem sido frequentemente referida como uma impossibilidade prática, traduzida por vezes, em expressões como comparar laranjas com maçãs (Iezzoni, 1996c), ou ainda de forma mais sistematizada, na impossibilidade de se compararem estádios de severidade no sistema de classificação Disease Staging (Plomman, 1982; Hornbrook, 1985 e Aronow, 1988).

Estes aspectos têm suscitado a dúvida sobre a utilização do diagnóstico como um factor de risco de morte.

Como foi referido, alguns sistemas defendem a criação de metodologias de medição da severidade específicas, ou seja por diagnóstico, como é o caso do Computerized Severity Index (CSI) (Horn et al, 1991). Outros defendem a aplicação de metodologias genéricas, como é o caso da grande maioria dos sistemas desenvolvidos para doentes tratados em Unidade de Cuidados Intensivos (APACHE e SAPS, por exemplo) (Knaus, Wagner e Draper, 1984; Knaus et al, 1985 e 1993; Le Gall, Lemeshow e Saulnier, 1993; Lemeshow et al, 1993). Existem ainda situações mistas, como é o caso do sistema Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatment (SUPPORT) (Murphy e Cluff, 1990), em que, pese embora o facto do sistema ser genérico na perspectiva da medição de cada uma das variáveis, é frequentemente utilizado na análise isolada por diagnóstico.

Comorbilidades

Por comorbilidades entende-se a existência de outros diagnósticos que não estão etiologicamente relacionados com o diagnóstico principal (Iezzoni, 1997d).

Na grande maioria das situações a estas comorbilidades está associada a ideia de doenças manifestadas pelos doentes no momento de admissão do hospital.

Estas comorbilidades têm sido frequentemente consideradas como indicadores válidos para se estabelecerem previsões sobre a morte a médio e longo prazo (Charlson et al, 1987), a curto prazo (Knaus et al, 1981; Hornbrook, 1982; Brewster et al, 1985; Blumberg, 1985; Knaus et al, 1985; Aronow, 1988; Horn, 1988; Philipps e Knaus, 1990; Greenfield et al, 1993; Le Gall, Lemeshow e Saulnier, 1993; Lemeshow et al, 1993; Ruttiman, 1994; Brailer et al, 1996; Elixhauser et al, 1998 e Gijssen et al, 2001) e inclusivamente, podem afectar a tolerância dos doentes para receberem determinado tratamento ou mesmo na respectiva resposta à terapêutica (Iezzoni, 1997d e Bleichrodt, Crainich e Eeckhoudt, 2003).

Situação/Estado funcional

A situação funcional é frequentemente considerada como outro dos factores de risco de morte. Isto porque, por exemplo, a dificuldade em andar pode resultar de outros factores, como anomalias congénitas, doenças neurológicas, etc... (Iezzoni, 1997d).

Neste sentido, a situação funcional respeita a áreas limitadas do funcionamento físico, mental e social nas actividades diárias. É, em síntese, o resultado do estado de saúde das pessoas (ausência de doença), do seu bem-estar (capacidade para participar activamente na “vida”) e da capacidade para suportar e ultrapassar os problemas de saúde (Rubenstein et al, 1989).

Contudo, atendendo a diversos problemas operacionais e conceptuais a situação funcional não tem sido muito utilizada para a definição de modelos de ajustamento pelo risco.

Outras dimensões de risco

As restantes dimensões do risco - a situação psicológica e cognitiva dos doentes, os atributos culturais, étnicos e socioeconómicos dos doentes e as atitudes e preferências dos consumidores, são igualmente considerados como relevantes no apuramento da probabilidade de morte (Kaplan, Greenfield e Ware, 1989; Applegate, Blass e Williams, 1990; Epstein et al, 1991; Iezzoni, 1991; Case et al, 1992; Bindman e Grumbach 1992 e Burstin, Lipsitz e Brennan, 1992).

Contudo, limitações de natureza prática, essencialmente as relacionadas com a dificuldade e fiabilidade na recolha dos dados têm conduzido a que não sejam dimensões muito utilizadas nos sistemas de ajustamento pelo risco.

2.3.3.3. Resultados de Saúde e Efectividade dos Cuidados Prestados

Conforme foi referido a escolha das características dos doentes para o modelo de ajustamento pelo risco depende essencialmente do tipo de resultado que se pretende analisar.

A título perfeitamente exemplificativo indicam-se alguns tipos de resultados (Iezzoni, 1995):

- Longevidade, morte;
- Estabilidade fisiológica;
- Complicações;
- Doenças crónicas e disfunções fisiológicas (morbilidade);
- Estatuto funcional, incapacidade;
- Desempenho funcional, desvantagem;
- Qualidade de vida;
- Custos de tratamento;
- Utilização de serviços (internamentos, readmissões, consultas);
- Demora Média.

Neste momento serão somente analisadas as questões relacionadas com a efectividade dos cuidados prestados, nomeadamente a mortalidade, as complicações e as readmissões (Thomas, Holloway e Guire, 1993; DesHarnais et al, 1997; Hill, Winfrey e Rudolph, 1997; HCIA, 1999; Chang, Lin e Northcott, 2002; NHS, 2002; Griffith, Alexander e Jelinek, 2002; O'Muircheartaigh, Murphy e Moore, 2002 e Selim et al, 2002).

Dentro destas deve-se ainda evidenciar que existem abordagens que representam resultados intermédios (complicações e readmissões) e um indicador de resultados finais – a mortalidade.

Em seguida irão ser analisadas as questões associadas com a relevância da utilização deste tipo de indicadores para a avaliação da efectividade e da qualidade dos hospitais.

Começando por discutir os aspectos respeitantes à mortalidade deve referir-se que a grande maioria dos autores considera que existe uma relação entre mortalidade e qualidade dos cuidados prestados (Blumberg, 1987; Dubois et al, 1987; Fink, Yano e Brook, 1989; lezzoni et al, 1992a; lezzoni et al, 1992b; lezzoni et al, 1992c; Krakauer et al, 1992; Landon et al, 1996; DesHarnais et al, 1997; Hill, Winfrey e Rudolph, 1997; Mennemeyer, Morrissey e Howard, 1997; HCIA, 1999; Chang, Lin e Northcott, 2002; Griffith, Alexander e Jelinek, 2002 e O'Muircheartaigh, Murphy e Moore, 2002).

Contudo, têm sido discutidos alguns aspectos, como a escolha do tipo de dados que deve ser utilizado e do sistema de classificação de doentes (lezzoni, 1993), sobre a janela de observação da mortalidade (Chassin et al, 1989; Garnick, DeLong e Luft, 1995 e Hughes et al, 1996) e se a mortalidade é ou não um resultado indesejável (lezzoni, 1995).

Na situação contrária, isto é, que não existe qualquer relação entre mortalidade e qualidade dos cuidados prestados encontram-se autores como Thomas e Hofer (1998 e 1999), os quais defendem que a taxa de mortalidade não é um bom indicador daquele atributo.

De facto Thomas e Hofer (1999) afirmam que a taxa de mortalidade ajustada pelo risco não é um bom indicador de qualidade, defendendo ainda que indicadores de processo são mais adequados para se medir a efectividade dos cuidados prestados.

Estes mesmos autores (Thomas e Hofer, 1998) afirmam que apesar de todos os esforços dedicados à análise da mortalidade hospitalar, tem sido dada pouca atenção à sua utilidade como indicador da qualidade dos cuidados prestados, podendo inclusivamente a publicação de dados deste tipo contribuir para a desinformação do público sobre o desempenho dos hospitais.

Ainda no mesmo artigo, estes autores referem que pese embora o facto de desde 1991 serem publicados dados nacionais (EUA) sobre o desempenho dos hospitais, sendo a taxa de mortalidade ajustada pelo risco o indicador mais utilizado, estas conclusões têm sido contestadas pelos hospitais, com o principal argumento de que dão indicações erróneas e enviesadas sobre a actividade hospitalar (Thomas e Hofer, 1998).

Para sustentarem as suas informações Thomas e Hofer (1998) formulam três hipóteses de investigação:

- Será que os hospitais que têm pior qualidade de cuidados prestados apresentam taxas de mortalidade ajustadas pelo risco mais elevadas?
- Podem os consumidores acreditar que os hospitais com taxas de mortalidade ajustadas pelo risco mais baixas são na realidade prestadores com qualidade?
- Para os hospitais será que as diferenças encontradas nas taxas de mortalidade ajustadas pelo risco estão correlacionadas com as diferenças observadas na qualidade prestada?

Em relação à primeira interrogação a resposta é positiva, enquanto que para as restantes a resposta encontrada pelos autores é negativa.

Assim, estes autores concluem que a sua revisão de literatura sugere que os doentes que receberam cuidados com qualidade mais baixa tiveram taxas de mortalidade mais elevadas. Encontraram igualmente que, em média, os doentes tratados em hospitais que foram identificados como “outliers” em termos de mortalidade elevada estão em maior risco de receberem cuidados de baixa qualidade. Contudo, algumas reservas e cautelas devem ser adoptadas quando se utiliza a mortalidade como indicador de qualidade.

Num estudo realizado por Bravo e outros (2002) pretendia-se analisar as consequências da regulação e da qualidade de cuidados na mortalidade no Canadá. As principais conclusões apontam para que a sobrevivência dos doentes não foi significativamente influenciada pelo tipo de regulação e que, pelo contrário, os doentes tratados com baixa qualidade apresentaram uma sobrevivência menor do que aqueles que tiveram acesso a cuidados de boa qualidade.

Por outro lado Fink, Yano e Brook (1989) referem que apesar das suas limitações a metodologia da Health Care Financing Administration (HCFA) tem assumido um papel importante na avaliação da qualidade dos cuidados prestados. Referem ainda que uma ligação directa entre a qualidade dos cuidados prestados e as taxas de mortalidade foi estabelecida por Dubois, Brook e Rogers (1987).

A este aspecto acrescem os argumentos apresentados por Iezzoni e outros, 1992b e Iezzoni, 1995, nos quais se defende que as principais vantagens da utilização de dados sobre a mortalidade hospitalar são a sua acessibilidade e a sua validade teórica como indicador da qualidade dos cuidados prestados e ainda que a grande maioria dos estudos sobre a efectividade consideram a mortalidade, essencialmente porque é facilmente definida e medida e apresenta

um valor claro para os doentes – na grande maioria das situações a morte é um resultado indesejável.

Em 1997 um estudo efectuado por Mennemeyer, Morrissey e Howard, embora não pondo em causa a utilização da taxa de mortalidade como indicador de qualidade, refere que os principais problemas na metodologia utilizada pela HCFA são os seguintes:

- Os dados dos hospitais apresentam erros e diferenças substanciais na codificação. Estes resultados podem condicionar os valores previsionais dos modelos (posteriormente esta temática será discutida);
- Os sistemas de classificação de doentes não medem com precisão a severidade do estado do doente;
- As medidas de severidade não distinguem entre problemas de saúde presentes no momento de admissão e aqueles adquiridos durante a hospitalização, tendo como principal consequência que os hospitais com pior qualidade parecem ter doentes mais graves;
- Como os hospitais têm diferentes práticas na admissão de doentes, tal pode conduzir a enviesamentos no tipo de doentes tratados e concomitantemente nos modelos de ajustamento pelo risco.

Convém ainda ter presente o defendido por Rosen e Green (1987) – as altas taxas de mortalidade não são um problema em si mesmo, devendo a principal preocupação residir na proporção de mortes evitáveis e ainda por Griffith, Alexander e Jelinek (2002), no qual é afirmado que os hospitais que apresentam melhores resultados, tanto na taxa de mortalidade ajustada pelo risco, como na taxa de complicações ajustada pelo risco, parecem demonstrar um melhor desempenho em termos de uma maior flexibilidade e de uma melhor adaptação às preferências dos consumidores.

Numa perspectiva diferente encontram-se os argumentos apresentados por Manheim e outros (1992), essencialmente quando referem que os principais resultados da investigação sugerem que a variação nas taxas de mortalidade são devidas a diferentes políticas e práticas dos hospitais na admissão de doentes, pelo que sem a existência de dados clínicos mais detalhados e acessíveis não é possível estabelecer uma relação determinística entre processo e resultados na prestação de cuidados hospitalares.

Green, Wintfeld e Passman (1990) defendem ainda que as taxas de mortalidade hospitalar podem ser úteis para os doentes e médicos se estes utilizarem estas indicações para escolherem o hospital que oferece a melhor perspectiva de sobrevivência. Contudo, referem igualmente que estes valores de mortalidade podem ser enviesados se a idade, a casuística e as necessidades de cuidados prolongados não forem consideradas para a avaliação do desempenho hospitalar.

Por outro lado, Jencks e outros (1988) referem ainda que a utilização de taxas de mortalidade ajustadas pelo risco poderão revelar grande utilidade para a avaliação da efectividade dos cuidados prestados, desde que a gravidade do doente seja devidamente avaliada no momento de admissão.

Outro aspecto complexo na definição de resultados em saúde respeita à janela de observação, isto é, ao período de tempo no qual o doente é seguido (Iezzoni, 1995). Neste contexto, para além de ser necessário distinguir entre resultados intermédios e resultados finais, é necessário que o período em análise seja indicado, pois podem ocorrer algumas situações em que a morte possa não ser um bom indicador de resultados.

Tudo isto tem suscitado um grande debate sobre o período de observação – mortalidade intra-hospitalar, mortalidade a 30 dias após a admissão ou mortalidade a 180 dias após a admissão (Iezzoni et al, 1995a)?

Em 1995, Garnick, DeLong e Luft concluem que a utilização de dados respeitantes à mortalidade a 180 dias após a admissão conduz a resultados semelhantes aos observados com a mortalidade a 30 dias, referindo ainda que estas conclusões são consistentes com o facto de a mortalidade ocorrida após 30 dias ter pouco a ver com as consequências de um episódio de internamento hospitalar.

Num estudo realizado em 1989, Chassin e outros encontram algumas diferenças entre resultados avaliados pela mortalidade intra-hospitalar e pela a mortalidade após 30 dias da admissão, afirmando que tal se pode dever essencialmente a diferentes práticas dos hospitais na alta dos doentes.

No campo contrário encontram-se os estudos de Knaus e outros (1985), Le Gall, Lemeshow e Saulnier (1993), Lemeshow e outros (1993) e Iezzoni (1997b e 1997c), nos quais se defende que a grande maioria dos sistemas de classificação de doentes contém algumas variáveis que são caracterizadoras e explicativas da mortalidade intra-hospitalar, embora não contenham as mesmas potencialidades quando se considera um prazo mais dilatado, a 30, 60 ou 180 dias. A este respeito Iezzoni e outros (1996a e 1996b) referem que apesar da consideração da mortalidade a 30 dias após a admissão ser importante, essencialmente derivado do facto de existirem políticas diferentes entre hospitais na alta dos doentes, não existem razões para acreditar, nem suficiente evidência científica que os resultados da avaliação dos hospitais em função da mortalidade seja diferente, quando se utiliza somente a mortalidade intra-hospitalar.

Atendendo a estes aspectos é de referir que essencialmente por razões operacionais, nomeadamente a disponibilidade e a fiabilidade dos dados que, a grande maioria dos estudos efectuados utilizam a mortalidade intra-hospitalar como medida de resultados de saúde. Este aspecto parece ainda ser mais pertinente em Portugal, visto que é praticamente impossível seguir a mortalidade

dos doentes a 30 ou a 180 dias, visto que estes dados não são recolhidos de uma forma consistente e rotineira pelas organizações de saúde.

Em termos de síntese, no que se refere à mortalidade pode-se acompanhar o defendido por Iezzoni (1993) – “Embora a credibilidade dos dados hospitalares fosse inicialmente contestada, vários estudos (Steen et al, 1993 e Krakauer e Jacoby, 1993) sugerem que a utilização de modelos de ajustamento pelo risco melhoram o valor predictivo, o rigor estatístico da análise” ou ainda por Krakauer e outros (1992) – “Atendendo a que o resultado do tratamento é fortemente dependente da doença e da severidade do estado do doente, a validade da inferência sobre o desempenho dos hospitais através da taxa de mortalidade ajustada pelo risco, está fortemente dependente da adequação do sistema de informação e do modelo utilizado para o ajustamento pelo risco”, pelo que para além de se ser pertinente utilizar a mortalidade como indicador de qualidade é igualmente necessário escolher um sistema de ajustamento pelo risco.

A relação entre os indicadores de resultados intermédios – complicações e readmissões – e a qualidade dos cuidados prestados não tem merecido tanta atenção internacional como a existente para a mortalidade.

Se por um lado se encontram autores em que esta associação é pacífica (Iezzoni et al, 1992d, 1994b, 1994c e 1997; DesHarnais et al, 1997 e 2000; HCIA, 1999 e Arozullah et al, 2003), também tem sido referenciado por outros (Thomas, 1996) que tal associação não está presente.

Este último autor não encontrou nenhuma evidência que sustentasse que os doentes que foram tratados com piores níveis de qualidade tenham apresentado uma maior probabilidade de serem readmitidos. Estes resultados de investigação são consistentes na grande maioria das 12 condições estudadas, tendo as excepções, ou seja os casos com níveis de readmissão distintos, sido encontradas para os casos médicos do foro cardíaco (readmissões a 60 dias) e para as doenças respiratórias (readmissões a 90 dias), mas com um sentido diferente ao esperado, ou seja, uma relação directa entre cuidados de qualidade e maiores readmissões (Thomas, 1996).

Num estudo realizado por Iezzoni e colegas em 1994b, sobre as complicações foram estabelecidas as seguintes hipóteses de investigação:

- Será que o desempenho relativo dos hospitais, medido pelas diferenças entre taxas de complicações observadas e esperadas é semelhante entre os serviços de cada hospital?

Os autores concluem que nos hospitais onde se observam maiores diferenças entre valores observados e esperados para um determinado serviço, também são aqueles onde se observa uma maior diferença na globalidade do hospital.

- Quais as características dos hospitais que estão associadas com as taxas de complicações?

O estudo refere que a dimensão, a função ensino e a oferta de serviços em cirurgia cárdio-torácica e de ressonância magnética são as características que conduzem a maiores diferenças entre valores observados e esperados.

- Estas características variam de serviço para serviço?

Os resultados apresentam comportamentos distintos entre casos médicos e cirúrgicos.

- Como é que se compara o desempenho dos hospitais quando se considera a taxa de complicações ou a taxa de mortalidade?

O estudo refere que o desempenho dos hospitais quando avaliado pelas complicações não está correlacionado com os valores da mortalidade, apresentando inclusivamente uma relação negativa nos casos médicos.

Num estudo realizado por Lawthers e outros (2000) tentou-se validar um instrumento, o “Complications Screening Program” (CSP), ao testar se as codificações derivadas da Classificação Internacional das Doenças, 9ª Revisão, Modificações Clínicas (CID-9-MC) são completas e precisas e ainda se o algoritmo do CSP permite identificar as condições presentes no momento de admissão daquelas adquiridas durante o tratamento. O CSP estabelece seis grupos de risco: grandes cirurgias, pequenas cirurgias, procedimentos cardíacos invasivos, endoscopia, casos médicos e total de casos. Os principais resultados referem que o CSP tem um bom desempenho na identificação de situações que podem ser consideradas como complicações, embora com comportamentos distintos entre casos cirúrgicos e médicos, com estes últimos a apresentarem piores valores. No que respeita à distinção entre complicações presentes na admissão e adquiridas durante o episódio de internamento a situação é problemática, sugerindo-se que a melhor solução passa pela criação de um código especial que permita identificar as situações presentes no momento de admissão.

Outro estudo realizado por Weingart e outros (2000) pretendia igualmente validar o CSP na identificação de complicações e de problemas potenciais na qualidade dos cuidados prestados. Os principais resultados referem que a probabilidade de ocorrer uma situação de baixa qualidade nas situações identificadas pelo CSP foram de 29.5% para os casos cirúrgicos e de 15.7% nos casos médicos. Pelo contrário, a probabilidade de encontrar uma situação com baixa qualidade nas situações não identificadas pelo CSP foi de 2.1%, tanto para casos cirúrgicos,

como médicos. Os autores, à semelhança do referido no estudo de Lamers et al (2000) concluem que o CSP apresenta desempenhos distintos para situações cirúrgicas e médicas, embora defendam que o CSP é eficiente para identificar problemas de qualidade, sugerindo ainda que mais investigação é necessária para se avaliar as percepções dos médicos sobre qualidade.

Estes resultados merecem uma série de reflexões suplementares, designadamente as associadas com os sistemas de informação existentes e com as diferentes práticas dos hospitais, tanto em termos de codificação como em termos de tratamento.

No que se refere aos sistemas de informação existentes deve referir-se que os mesmos não disponibilizam dados sobre as complicações presentes no momento de admissão e as adquiridas durante o processo de tratamento. Este facto, só por si, pode justificar que existam situações em que doentes com mais complicações adquiridas durante o processo de tratamento apresentem resultados diferentes em termos de mortalidade, devido a diferenciais das condições pré-existentes (Iezzoni, 1997b).

As diferentes práticas na codificação, com as necessárias repercussões nos diagnósticos secundários e concomitantemente nas complicações, podem levar a que hospitais apresentem piores relações entre valores esperados e observados nas complicações, embora o mesmo não tenha forçosamente de se passar quando se analisa a mortalidade (Kalish et al, 1995 e Weingart et al, 2000).

No que se refere às práticas de tratamento dos hospitais pode acontecer que alguns hospitais tratem os doentes com maior agressividade e que conseqüentemente aumentem as complicações, embora simultaneamente se observe uma melhor qualidade e uma diminuição das mortes (Weingart et al, 2000).

Neste sentido deve evidenciar-se que, com excepção dos estudos realizados por Thomas (Thomas 1996 e Thomas e Hofer, 1998 e 1999), praticamente ninguém contesta a relação entre indicadores de resultados, intermédios e finais, com a efectividade dos cuidados, chamando no entanto a atenção para a necessidade de se definirem modelos de ajustamento pelo risco e ainda para algumas dificuldades conceptuais e operacionais para se atingir este objectivo.

2.3.3.4. Resultados de Saúde e Eficiência dos Cuidados Prestados

A eficiência técnica é uma medida que traduz uma relação entre os recursos utilizados e os resultados obtidos (Pereira, 1993).

Tradicionalmente os indicadores mais utilizados para avaliação da eficiência técnica dos hospitais são os custos médios e a demora média (Evans, 1981; McGuire, Henderson e Money, 1988 e Butler, 1995).

Existe igualmente uma grande discussão a nível internacional sobre a validade e comparabilidade destes dois indicadores, sendo de evidenciar algumas diferenças entre os valores teóricos e os resultados de estudos empíricos.

Na realidade, atendendo essencialmente à intensidade dos recursos admite-se que custos e duração de internamento tenham comportamentos distintos, visto que, por um lado, a intensidade dos recursos diminui à medida que aumenta a duração de internamento e que, por outro lado, os custos aumentam à medida que diminui a duração de internamento (Butler, 1995 e Evans, Hwang e Nagarajan, 1995).

Por intensidade na utilização de recursos entende-se a quantidade de recursos utilizados num determinado período de tempo, por exemplo por dia de internamento (Luke, 1979).

No entanto, vários estudos empíricos têm encontrado uma associação forte entre custos ou “charges” e a demora média (Lave e Leinhardt, 1976; Jones, 1985; Coffey e Goldfarb, 1986; Horn et al, 1991; Knaus et al, 1993; Bradbury, Golec e Steen, 1994; Chen e Naylor, 1994; Mozes et al, 1994; Becker et al, 1995; MEDSTAT, 2001; Clark e Ryan, 2002 e Polverejan et al, 2003).

Por outro lado, embora existam diversas críticas à utilização da duração de internamento, dentro das quais avultam as questões da intensidade de utilização de recursos e a prática distinta dos hospitais em relação à alta dos doentes (Knaus et al, 1993; Costa, 1994 e Butler, 1995), existem igualmente argumentos contra a utilização dos custos médios, designadamente os relacionados com as diferentes práticas contabilísticas dos hospitais (Ament et al, 1982; Berki, Ashcraft e Newbrander, 1984; Brewster et al, 1985; Hornbrook e Monheit, 1985; Thomas, Ashcraft e Zimmerman, 1986 e Vladeck et al, 1988).

A este propósito Mozes et al (1994) referem inclusivamente que a utilização da demora média é preferível à dos custos, visto que estes são muito mais sensíveis a alterações nos procedimentos de apuramento, tanto em termos históricos, como geográficos.

Quando se considera a realidade portuguesa esta polémica é mais fácil de ultrapassar, dado que na sua grande maioria os hospitais portugueses apuram custos por serviço, utilizando ainda critérios de distribuição e imputação de rubricas de despesa distintos.

Neste sentido, de seguida serão discutidos os aspectos associados com a duração de internamento ou com a demora média em relação:

- Ao “rationale” ou suporte teórico do ajustamento pelo risco;
- À evidência prática do ajustamento pelo risco para avaliação da demora média;

- À pesquisa de eventuais contradições entre a efectividade dos cuidados prestados e a eficiência técnica.

A discussão sobre a eficiência dos hospitais e mais concretamente sobre a demora média ganhou maior relevo após a introdução do pagamento prospectivo por caso nos EUA em 1983 (Iezzoni et al, 1996c), sendo frequentemente discutido se a gravidade dos doentes afecta ou não o consumo de recursos (Vladeck, 1984; Coffey e Goldfarb, 1986; Jencks e Dobson, 1987; Lave e Franck, 1990; Horn et al, 1991; Burns, Wholey e Abeln, 1993; Kelleher, 1993; Chen e Naylor, 1994; Mozes et al, 1994; Voss et al, 1994; Schwartz et al, 1996; Marazzi et al, 1998 e Clark e Ryan, 2002).

Praticamente todos os estudos anteriormente referidos apontam para grandes discrepâncias na demora média entre os hospitais, sendo frequentemente referido que as principais razões justificativas podem ser encontradas nas características dos doentes (idade, doença, gravidade e situação sócio-económica), nas características dos hospitais (dimensão, desenvolvimento tecnológico, função ensino e preferências de tratamento dos prestadores) e nas condições do mercado concorrencial (localização, extensão da oferta e modalidades de pagamento) (Goldfarb, Hornbrook e Higgins, 1983; Coffey e Goldfarb, 1986; Lave e Franck, 1990; Thomas et al, 1998 e Rapoport et al, 2003).

Na sua grande maioria os estudos apontam para a necessidade de se ajustar pelo risco, controlando desta forma algumas das características dos doentes, para melhor se comparar a demora média entre os hospitais, sendo ainda referido que os sistemas de classificação de doentes que apresentam um bom desempenho para prever a mortalidade podem não ter um resultado idêntico para estabelecerem previsões sobre custos ou duração de internamento (Voss et al, 1994).

Tal pode ser devido a questões estruturais, como a existência de contradições entre gravidade e consumo de recursos, bem patente na situação dos mortos ou dos cuidados paliativos (Coffey e Goldfarb, 1986; Murphy e Cluff, 1990 e Knaus et al, 1993) ou ainda a diferentes relações entre as mesmas variáveis para explicar a mortalidade ou o consumo de recursos (Goldfarb, Hornbrook e Higgins, 1983; Schwartz et al, 1996; Iezzoni et al, 1996c; MEDSTAT, 2001 e Houchens, 2002).

Num estudo realizado por Schwartz e outros (1996) para episódios de internamento com fractura da anca, os autores pretenderam investigar a importância da gravidade para explicar variações na duração de internamento ao nível de cada doente e ainda a contribuição das medidas de gravidade para explicar diferenças na demora média entre hospitais, tendo na generalidade chegado às seguintes conclusões:

- A gravidade tem um poder explicativo reduzido ao nível da duração de internamento por doente;
- Praticamente todos os sistemas de classificação de doentes utilizados (14 sistemas diferentes) apresentam desempenhos semelhantes na explicação individual da duração de internamento, embora o valor explicativo medido pelo coeficiente de determinação seja mais elevado com os All Patient Refined Diagnosis Related Groups (APRDRGs) e com o Disease Staging (numa versão que inclui as comorbilidades desenvolvida por Naessens et al, 1992);
- A gravidade tem igualmente um bom poder explicativo quando se analisam as diferenças na demora média entre hospitais, tendo ainda todos os sistemas de classificação de doentes apresentado uma boa concordância na ordenação dos hospitais;
- Referem também que o ajustamento pela gravidade revelou pouca importância na ordenação dos hospitais, visto que esta não é significativamente diferente da obtida quando se utiliza a demora média;
- No entanto, afirmam que apesar dos resultados encontrados o ajustamento pelo risco revela alguma utilidade para a revisão de utilização e para se melhorar o desempenho na duração do internamento a nível individual, visto que as diferenças encontradas na demora média entre estadios de gravidade são importantes (entre 50% a 100% superiores nos doentes mais graves).

Outro estudo realizado em 1996 por Iezzoni e outros (1996c), pretende estudar as mesmas questões, mas agora com aplicação aos episódios de internamento com pneumonia. Na generalidade, as conclusões são idênticas às referidas no estudo anterior, o que pode ser considerado normal, visto que os autores são os mesmos.

Num estudo realizado por Becker e outros (1995) estudou-se a importância da gravidade (medida pelo APACHE II) para avaliar o desempenho das Unidades de Cuidados Intensivos no que se refere à demora média, consumo de recursos e mortalidade nos doentes que tiveram uma intervenção cirúrgica cardíaca.

Neste estudo o ajustamento pelo risco demonstrou um bom poder explicativo sobre a demora média no que se refere a grupos de doentes, embora com poder explicativo reduzido ao nível individual (Becker et al, 1995).

No entanto, referem ainda que estes resultados são normais, visto que o ajustamento pelo risco não se destina a aplicações individuais (doente a doente), visto que acreditam que a comparação entre valores observados e esperados para a demora média para grupos de doentes, permite aos médicos melhorar a eficiência dos cuidados prestados (Becker et al, 1995).

Num estudo realizado por Knaus e outros (1993) sobre as variações na demora média e na mortalidade em Unidades de Cuidados Intensivos, os autores

referem a importância do ajustamento pelo risco para se avaliarem os resultados.

Referem ainda que a duração de internamento é frequentemente utilizada como uma medida dos custos, apesar dos problemas decorrentes da diferente intensidade de cuidados prestados e que a contribuição da gravidade é mais relevante para explicar diferenças na mortalidade e que para a duração de internamento devem ser consideradas conjuntamente a gravidade e a doença principal para se aumentar o respectivo poder explicativo (Knaus et al, 1993).

Finalmente estes autores referem ainda a existência de potenciais problemas entre a gravidade e a demora média, essencialmente pelo diferente comportamento apresentado pelos doentes falecidos, recomendando ainda que se devem realizar estudos mais aprofundados sobre a duração de internamento, utilizando interações entre doenças e gravidade (Knaus et al, 1993).

Num estudo realizado por Bradbury, Golec e Steen (1994) foi encontrada uma associação positiva entre a mortalidade ajustada pelo risco e os custos ajustados pelo risco e a demora média ajustada pelo risco, embora com significância estatística somente para a associação mortalidade/demora média. Neste estudo é igualmente referido que as características dos hospitais analisadas (dimensão, recursos humanos, estatuto universitário e localização em áreas urbanas) não estão relacionadas com a demora média.

Num estudo realizado por Coffey e Goldfarb (1986) analisaram-se as diferenças entre os DRGs e o Disease Staging para analisar a demora média e o financiamento dos hospitais, sendo de evidenciar os seguintes aspectos:

- Os dois sistemas de classificação de doentes são conceptual e empiricamente diferentes, nomeadamente no facto de os DRGs classificarem doentes por procedimentos, enquanto que tal não ocorre no Disease Staging;
- Apesar destas diferenças entre os dois sistemas de classificação de doentes, os mesmos apresentam um desempenho semelhante para explicar as variações na demora média entre hospitais;
- No entanto, existem diferenças significativas nos pagamentos aos hospitais, com as organizações de saúde com maior dimensão a serem beneficiadas pelos DRGs;
- Referem ainda que os hospitais com maior dimensão tratam doentes menos graves, embora com Índices de Casemix dos DRGs mais elevados, essencialmente porque estas organizações aceitam maiores proporções de doentes menos graves para optimizarem a sua ocupação.

Num estudo realizado por Westert e outros (1993) analisou-se a variação na demora média entre hospitais e entre médicos no mesmo hospital. Após ajustar

a demora média pelo risco os autores referem que, por doença, as diferenças encontradas na demora média por médicos no mesmo hospital, são insignificantes, quando comparadas com as diferenças na demora média encontradas entre hospitais e que existe uma grande variação na prática médica, quando o mesmo médico trabalha em dois hospitais diferentes.

Num estudo realizado por Chen e Naylor (1994) analisou-se a variação na demora média dos hospitais para o enfarte agudo do miocárdio. Os autores referem que existe uma grande variação na demora média entre os hospitais, a qual é no entanto reduzida quando se retiram os óbitos, as transferências e os doentes com revascularização. Referem ainda que a duração de internamento é uma componente importante dos custos por doente e que as características dos doentes, designadamente a gravidade, contribuem para explicar 12.6% da demora média. Finalmente referem que as características dos hospitais não revelam utilidade para prever a demora média.

Num estudo realizado por Voss e outros (1994) é referido que a gravidade medida por critérios clínicos produziu melhores previsões para a demora média que os DRGs, embora uma parte substancial na variação do consumo de recursos fique por explicar. Neste estudo é ainda evidenciado que o ajustamento pelo risco para se avaliar a demora média é mais relevante para os casos médicos, em comparação com os casos cirúrgicos.

Deve ainda referir-se que existe uma variedade de outros estudos (Jones, 1985; Lave e Franck, 1990; Cleary et al, 1991; Horn et al, 1991; Burns, Wholey e Abeln, 1993; Kelleher, 1993; Mozes et al, 1994; Yuan et al, 2000; Kaboli et al, 2001 e Clark e Ryan, 2002), os quais, pese embora o facto de terem sido aplicados a populações distintas, evidenciam praticamente os mesmos resultados, nomeadamente que a gravidade aumenta o poder justificativo do consumo de recursos e que o ajustamento pelo risco é útil para avaliar a eficiência dos hospitais.

Finalmente, devem ser discutidas as eventuais contradições entre a optimização da efectividade dos cuidados prestados e a optimização da respectiva eficiência.

Bradbury, Golec e Steen (1994) começam por situar o problema. Por um lado, face à teoria da produção (input/output) é de esperar uma relação negativa, ou seja, quanto mais recursos aplicados menor é a mortalidade, no entanto, por outro lado, referem outro tipo de argumentos que contrariam esta relação, designadamente o facto de a melhoria dos resultados de saúde estar associada a uma maior eficiência na utilização de recursos (por exemplo, a diminuição das complicações pode implicar uma redução nos custos) e de uma maior efectividade no diagnóstico e tratamento dos doentes implicar igualmente uma redução no consumo de recursos.

No referido estudo os autores acabam por concluir que existe uma associação positiva entre consumo de recursos e a mortalidade, embora somente com significância quando se utiliza a duração de internamento para se medir o consumo de recursos (Bradbury, Golec e Steen, 1994).

Outros estudos internacionais apresentam resultados diferentes. Enquanto que Stearns (1991) refere que não existem contradições entre a mortalidade e o consumo de recursos, Chen e Naylor (1994) afirmam que não existe qualquer relação entre a mortalidade e a demora média.

Por outro lado, Voss e outros (1994) e Yuan e outros (2000) apresentam diferenças no comportamento dos hospitais para optimizarem a efectividade e a eficiência, enquanto que Knaus e outros (1993) e Dismuke e Sena (1998) vão ainda mais longe, visto que avançam com contradições na prossecução dos dois objectivos, embora os primeiros avancem com argumentos teóricos e os segundos apresentam resultados empíricos de hospitais portugueses.

Na mesma linha de pensamento, ou seja comportamento distinto entre a gravidade e a demora média, surgem os artigos de lezzoni e outros (1996c) e Shwartz e outros (1996), os quais utilizem a probabilidade de morte e a probabilidade de morte ao quadrado, para pesquisar se a duração de internamento é menor em estadios de gravidade menos e mais elevados.

Em síntese, pode afirmar-se que subsiste ainda uma grande polémica sobre a eventual existência de contradições para se optimizar a efectividade e a eficiência dos hospitais, tanto à luz dos comportamentos teóricos esperados, como em função dos resultados empíricos, pelo que se torna um assunto com necessidade de aprofundamento.

2.3.3.5. Modelos de Ajustamento pelo Risco

A recolha de dados para o desenvolvimento de um sistema de severidade e consequentemente de ajustamento pelo risco apresenta, na generalidade, dois tipos de preocupações:

- A definição dos suportes de recolha da informação;
- A definição dos momentos de medição.

Existe uma grande discussão sobre os suportes de recolha da informação que devem ser utilizados para o desenvolvimento de um sistema de medição da severidade.

Na realidade, enquanto que alguns autores (Fetter et al, 1980; Young, Swinkola e Zorn, 1982; Gonnella, Hornbrook e Louis, 1984; Gonnella et al, 1990 e Young, Kohler e Kowalsky, 1994, por exemplo) defendem a utilização dos resumos

informatizados de alta, baseados no Uniforme Hospital Discharge Data Abstract (UHDDA), outros (Horn, 1986 e 1988; Knaus et al, 1981, 1985 e 1986; Brewster et al, 1995; Le Gall, Lemeshow e Saulnier, 1993; Lemeshow et al, 1993 e Moreno, Apolone e Miranda, 1998, por exemplo) argumentam que os elementos dos resumos informatizados de alta devem ser complementados com outros presentes nos processos clínicos dos doentes.

As principais razões para estas discrepâncias podem ser sintetizadas pelo seguinte balanço: privilegiar a operacionalidade e economia dos sistemas de classificação de doentes ou, privilegiar a validade de construção e de conteúdo dos sistemas de classificação de doentes.

Na validade de construção está-se a determinar a correspondência entre os valores apresentados pelo sistema e os valores reais, ou seja, será que o sistema está a medir aquilo que pretende (Hornbrook, 1982 e Thomas, Ashcraft e Zimmerman, 1986).

E, na validade de conteúdo está-se a apurar a lógica interna dos sistemas para definir categorias e o seu grau de credibilidade e de compreensão por parte dos profissionais de saúde (Hornbrook, 1982 e Thomas, Ashcraft e Zimmerman, 1986).

Os princípios de operacionalidade e de economia, embora se encontrem aplicados a planos conceptuais diferentes, apelam para a maior acessibilidade e disponibilidade dos dados proporcionada pelos resumos informatizados de alta.

Por outro lado, a necessidade de recurso suplementar a dados constantes nos processos clínicos dos doentes, para além de dificultar essa mesma acessibilidade e disponibilidade, apela para uma maior compreensão e extensão dos dados recolhidos, o que implicará uma maior consistência e validade dos próprios sistemas de classificação.

Esta distinção tem sido crucial para o desenvolvimento dos próprios sistemas de classificação. Na realidade os sistemas que utilizam preferencialmente os resumos de alta definem e medem risco em termos da quantidade dos órgãos afectados. Pelo contrário, os sistemas que suplementarmente recorrem aos elementos dos processos clínicos definem e medem esse risco não pela quantidade dos órgãos afectados, mas pela extensão da gravidade das insuficiências de um ou mais órgãos.

Em última instância o dilema que se coloca reside na definição de prioridades e daquilo que se pretende sacrificar.

Quando se entende que os aspectos financeiros são determinantes, então será natural que se privilegie o recurso quase exclusivo a elementos dos resumos de alta. Quando se defende que a validade de construção e de conteúdo de um sistema de classificação de doentes para medição da severidade é um aspecto a preservar, então o recurso aos elementos dos processos clínicos é inevitável.

Um outro plano que simultaneamente deverá ser considerado para a recolha de dados diz respeito à fiabilidade. Em relação ao conceito fiabilidade estão associados dois tipos de questões: a possibilidade de o sistema ser utilizado da mesma forma por parte de diferentes classificadores ou pelo mesmo classificador em diferentes períodos de tempo; e a possibilidade de o sistema proporcionar aos hospitais a classificação de doentes em situações que lhes são mais favoráveis, para efeitos de financiamento ou para otimizar resultados, por exemplo (Hornbrook, 1982).

Estas questões, bem como outras relacionadas com as próprias medidas de fiabilidade são discutidas num artigo de Costa e Nogueira (1994). As principais conclusões referem o facto de os sistemas de classificação que utilizam os resumos de alta (basicamente informação sobre diagnóstico principal e diagnósticos secundários) apresentarem na generalidade níveis de fiabilidade mais baixos que os sistemas que recorrem aos processo clínicos (basicamente informação sobre sinais vitais e resultados de exames complementares de diagnóstico).

As principais razões estão associadas ao maior grau de subjectividade que os sistemas baseados na identificação de diagnóstico apresentam, inclusivamente devido à sobreposição de categorias existente no CID-9-MC (McMahon e Smits, 1986), sistema utilizado para codificação dos diagnósticos nos hospitais, em oposição à objectividade dos elementos proporcionados pelos sinais vitais e resultados dos exames complementares de diagnóstico em que praticamente os problemas que podem existir dizem respeito somente ou a perdas de informação ou a erros de transcrição, o que confere uma maior objectividade aos sistemas que recorrem a estes elementos (Iezzoni et al, 1995b; Iezzoni, 1997b; Iezzoni et al, 1996d e Iezzoni et al, 1996e).

A definição dos momentos de medição dos dados tem repercussões em dois níveis de análise:

- Na neutralidade económica do sistema;
- Na prospectividade do sistema.

Com a neutralidade económica discute-se se o sistema de classificação deve ou não deve ser indiferente em relação ao desempenho dos hospitais (Jencks et al, 1984). Ou seja, se a multiplicidade de tratamentos que existem ou estão disponíveis em determinado hospital deve ou não ser considerado como mais um factor para se desagregar ou definir novos produtos. Por exemplo, um sistema com neutralidade económica considera produtos diferentes situações clínicas com o mesmo diagnóstico e prognóstico mas com tratamento médico ou cirúrgico.

Quando se defende que um sistema de classificação deve ter neutralidade económica, então na sua definição e operacionalização está implícita a

incorporação da eficiência de tratamento, da qualidade dos cuidados prestados, da disponibilidade de “inputs”, bem como da natureza e manifestação da doença (Costa, 1994).

Por outro lado, resulta lógico que para a obtenção de um sistema que tenha neutralidade económica devem ser considerados elementos respeitantes a todos os momentos do episódio de internamento. Ou seja, elementos representantes da admissão, da estadia (máximos ou médios) e da alta.

Contudo, tem frequentemente sido referido (Thomas et al, 1998; Thomas e Longo, 1990; Iezzoni et al, 1988, 1991a e 1991b e Greenfield et al, 1988) que a inclusão de elementos respeitantes a todo o episódio de internamento, para além de diminuir o risco da actividade das organizações de saúde, pode originar problemas suplementares derivados do registo de doenças iatrogénicas, o que, em última instância, pode limitar a fiabilidade e a validade do sistema de classificação, já que idênticas situações de saúde, à partida, podem conduzir a produtos distintos, dependendo da disponibilidade de “inputs”, da eficiência e da efectividade dos próprios hospitais.

A prospectividade do sistema de classificação constitui um factor relativamente fácil de descrever. A mesma só poderá ser assegurada quando um sistema se baseia nos elementos de admissão, visto que a consideração de elementos da estadia e ou da alta condicionam a respectiva classificação para um período imediatamente posterior à alta do doente (Green, Passman e Wintfeld, 1991).

Atendendo a que no processo de decisão clínica existe a necessidade de decidir a “quente”, ou seja imediatamente a seguir à admissão dos doentes, resulta natural admitir que um sistema de classificação prospectivo será aquele que sintetiza melhor o processo de produção nos hospitais.

Dos argumentos atrás referidos fácil será inferir que um sistema de classificação ideal não deve ter neutralidade económica e deve ser prospectivo.

A forma de contacto com as organizações de saúde, designadamente a explicitação da dicotomia entre admissões programadas ou urgentes, bem como o tipo de tratamento, expresso pela identificação de tratamentos médicos ou cirúrgicos, têm sido igualmente apontados como factores que podem justificar riscos de morte distintos.

Na realidade, sistemas como o APACHE III ou o SAPS, por exemplo, têm sustentado que a forma de admissão (programada ou urgente) é um factor predictivo de morte (Knaus et al, 1985; Le Gall, Lemeshow e Saulnier, 1993; Carneiro, 1994), com riscos superiores nas admissões urgentes, essencialmente devido ao facto dos doentes admitidos por esta via estarem menos estudados e/ou eventualmente, serem mais graves.

Esta mesma conclusão foi retirada por Costa, Delgado e Carvalho (1989) num estudo efectuado num hospital português, o qual para além de identificar níveis de severidade mais elevados em doentes admitidos pela urgência, referia igualmente um maior consumo de recursos neste tipo de doentes.

Para o tipo de tratamento, médico ou cirúrgico, existe alguma diversidade nos sistemas desenvolvidos até ao presente momento.

Sistemas como o APACHE I, CSI, Disease Staging, MPM e MedisGroups, por exemplo, não consideram que a presença de intervenção cirúrgica seja um factor adicional explicativo da morte (Iezzoni, 1997b).

Pelo contrário os sistemas APACHE II e III, DRGs, AIM, SAPS e PMCs, por exemplo, entendem que a presença de intervenção cirúrgica é mais um factor explicativo da morte (Iezzoni, 1997b).

Atendendo a esta contradição, bem como ao facto de a presença ou ausência de intervenção cirúrgica constituir um dos factores que potencialmente pode explicar níveis distintos de utilização de recursos (Hornbrook, 1982 e 1985; Hornbrook e Monheit, 1985; McMahon e Billi, 1988; Ahicart, 1988; Charbonneau et al, 1988 e Rosko, 1988), a presença/ausência de tratamento cirúrgico poderá constituir um elemento importante para a criação de modelos de ajustamento pelo risco.

Os principais problemas referentes à explicitação final do sistema de classificação dizem respeito a dois aspectos (Knaus et al, 1986; Ruttiman, 1994 e Iezzoni, 1997c):

- Ao tipo de escala apresentado pelo sistema de classificação;
- À formulação final do sistema de medição.

Existem quatro tipos de escalas, de níveis hierarquizados (Melo, 1985): nominais, ordinais, de intervalos e de razões.

Segundo o mesmo autor (Melo, 1985):

- As escalas nominais são constituídas por um conjunto de classes de equivalência, sendo esta estabelecida por um critério empírico ligado à realidade a “medir”;
- As escalas ordinais estabelecem uma relação de maior ou menor em relação às categorias adjacentes, não permitindo relacionar aritmeticamente as diferenças entre as diversas categorias;
- As escalas de intervalos são aquelas em que é definida uma igualdade de intervalos;

- As escalas de razão permitem a comparação, não só da amplitude de intervalos, mas também da respectiva razão.

Esta distinção não é meramente estatística, visto que para efeitos de gestão ou de financiamento, por exemplo, interessa não só identificar produtos (definir categorias), mas também comparar esses mesmos produtos/categorias.

Com escalas nominais ou ordinais, somente a primeira finalidade (identificação de produtos) poderá ser atingida, necessitando-se no mínimo de uma escala de intervalos para se proceder às posteriores comparações e avaliações.

Esta problemática tem sido evidenciada em quase todos os sistemas desenvolvidos até ao momento.

Na realidade, os sistemas DRGs e PMCs apresentam escalas do tipo nominal, isto é, identificam somente categorias. Os sistemas AIM, Disease Staging, CSI e MedisGroups, apresentam escalas do tipo ordinal. Os sistemas APACHE, MPM, SAPS, apresentam escalas de intervalos ou de razão. Por outro lado algumas versões do Disease Staging, dos PMCs, ou dos MedisGroups, por exemplo, apresentam escalas de intervalos. O mesmo acontece quando os DRGs são aplicados para efeitos de financiamento, em que a atribuição de pesos relativos para este efeito, confere à escala potencialidades do tipo de escala de razão (Iezzoni, 1997b).

Do atrás referido, resulta que se verificam preocupações crescentes para a adopção de sistemas de classificação que apresentem, no mínimo, escalas de intervalos.

Contudo, atendendo a que a maior compreensão e aceitabilidade por parte dos médicos constitui igualmente um aspecto da maior importância, será natural que noutra fase se possa igualmente apurar escalas distintas para cada variável considerada no modelo final (Lemeshow et al, 1993; Ruttiman, 1994).

Mas será que estas questões conceptuais e operacionais sobre a definição de sistemas de classificação de doentes para ajustamento pelo risco têm implicações na avaliação da efectividade e da eficiência dos hospitais?

A resposta a esta questão não é pacífica, visto que, como anteriormente foi referido, existem autores que defendem a utilização de modelos baseados nos resumos de alta, enquanto que outros defendem a supremacia dos modelos baseados nos dados dos processos clínicos, para finalmente outros argumentarem que a utilização de uns ou outros é indiferente, pelo que o processo de decisão sobre a escolha de determinado sistema de classificação de doentes deve obedecer a critérios mais pragmáticos, como a sua exequibilidade e os respectivos custos de implementação e de exploração.

Na realidade Krakauer e outros (1992), embora começando por afirmar que a avaliação dos resultados do tratamento dos hospitais seja fortemente dependente da “doença”, do doente e da sua gravidade, a validade da inferência sobre o desempenho dos hospitais através da mortalidade ajustada pelo risco está fortemente dependente da adequação do sistema de informação e do sistema utilizado para ajustamento pelo risco, acabam por concluir que existe uma concordância significativa entre a mortalidade prevista, tanto quanto se utilizam dados dos resumos de alta, como os respeitantes aos processos clínicos.

Isto porque nos estudos que efectuaram, constataram que os modelos de ajustamento pelo risco apresentam valores moderados para a sensibilidade (81%) e para a especificidade (79%). Por outro lado, o seu valor predictivo negativo é alto (90%), pelo que concluem que se um hospital não é “outlier” no modelo administrativo, também é provável que o não seja no modelo clínico. Contudo, o valor predictivo positivo é baixo (64%), pelo que afirmam que se um hospital é “outlier” no modelo administrativo, existe ainda uma probabilidade forte de o não ser no modelo clínico (36%) (para melhor perceber os conceitos estatísticos referidos, ver capítulo seguinte).

No entanto, atendendo a que os resultados de investigação parecem demonstrar que a classificação dos hospitais é muito mais estável quando se utilizam modelos baseados nos dados administrativos, os autores acabam por concluir que estes modelos de ajustamento pelo risco contêm informação suficiente para se avaliarem as variações nos resultados de saúde e consequentemente no desempenho dos hospitais.

Na mesma linha de pensamento encontram-se os argumentos de DesHarnais e colegas (2000), os quais para além de referirem as vantagens económicas da utilização de modelos de ajustamento pelo risco baseados nos resumos de alta, referem ainda que os resultados de aplicação são, no mínimo, similares aos disponibilizados pelos modelos que recorrem aos processos clínicos.

Numa perspectiva distinta pode referir-se o estudo de Hughes e outros (1996), no qual os autores defendem que a avaliação da precisão do modelo de ajustamento pelo risco, para além de depender do tipo de risco que está a ser avaliado, poderá ainda ser influenciado por algumas perspectivas contraditórias. A este respeito, os autores afirmam que embora se espere que o consumo de recursos, a duração de internamento e o risco de morte apresentem uma associação forte, a existência de um volume importante de mortes prematuras pode contrariar este comportamento esperado.

Numa atitude mais céptica encontra-se o estudo de Manheim e outros (1992), os quais defendem que a variação encontrada nas taxas de mortalidade ajustadas pelo risco são devidas a diferentes políticas e práticas dos hospitais na admissão de doentes. Assim, acabam por concluir que sem a existência de dados clínicos

mais detalhados e acessíveis, não é possível estabelecer uma relação determinística entre o processo e a mortalidade.

Este tipo de argumento vem na sequência de as conclusões retiradas por Greenfield e outros (1988) e de Green, Wintfeld e Passman (1990). No primeiro artigo, os autores concluem que as comorbilidades devem ser consideradas em qualquer metodologia de avaliação da qualidade, pelo que se tal não for considerado, as variações nas práticas e nas políticas de admissão de cada hospital podem contribuir para interpretações erróneas sobre a qualidade hospitalar. O estudo de Green, Wintfeld e Passman (1990) refere a necessidade de a taxa de mortalidade hospitalar ser corrigida pela severidade do estado do doente presente no momento de admissão, para desta forma se estabelecerem inferências válidas sobre a qualidade ou efectividade dos cuidados prestados.

Num artigo publicado por Iezzoni (1997c) é descrito um estudo na qual são comparados os desempenhos estatísticos de cinco sistemas de classificação – MedisGroups, APACHE (“score” fisiológico), Disease Staging, Patient Management Categories e All Patient Refined – Diagnosis Related Groups (APRDRGs) – são utilizados dois indicadores para avaliar o desempenho estatístico, a estatística “c” e o coeficiente de determinação (r^2), tendo chegado à conclusão que as medidas baseadas nos resumos clínicos de alta (Disease Staging, Patient Management Categories e All Patient Refined Diagnosis Related Groups) apresentam, na generalidade melhores resultados.

No mesmo artigo e quando se comparam as taxas de mortalidade esperadas em cada modelo, é referido que os sistemas de classificação utilizados apresentam resultados diferentes, com maiores discrepâncias entre medidas baseadas nos resumos de alta e as que utilizam os processos clínicos. Contudo, tal também está presente entre medidas baseadas somente nos resumos de alta (PMCs e APRDRGs, por exemplo) e nos processos clínicos (entre o APACHE e os MedisGroups, para as pneumonias).

Neste sentido Iezzoni (1997c) acaba por concluir que os hospitais devem estar preocupados com as comparações entre as taxas de mortalidade observadas e as esperadas, visto que os resultados são diferentes em função do sistema de classificação de doentes utilizado.

Refere ainda que, apesar do seu poder explicativo, a utilização de sistemas de classificação de doentes baseados nos resumos informatizados de alta, pode suscitar alguns problemas para se avaliar a qualidade, essencialmente porque estas medidas estão fortemente dependentes dos problemas decorrentes da CID-9-CM (McMahon e Smits, 1986), do nível de codificação (Jencks, Williams e Kay, 1988; Iezzoni e Moskowitz, 1986; Iezzoni et al, 1992d e Green e Wintfeld, 1993) e dos incentivos financeiros para se proceder à codificação (Simborg, 1981).

Finalmente afirma que, sempre que possível, deve ser feita uma distinção entre as condições presentes no momento da admissão e as adquiridas durante o internamento, para que desta forma se evite o “death code creep”, como forma de otimizar os resultados de saúde de cada hospital (Iezzoni, 1997c).

Noutro estudo realizado por Iezzoni e outros (1995a), os autores pretendem comparar a qualidade dos hospitais, medida pela mortalidade, em função de diversos sistemas de classificação de doentes. Os modelos de ajustamento pelo risco utilizados foram os MedisGroups, APACHE (“score” fisiológico), Disease Staging e All Patient Refined – Diagnosis Related Groups (APRDRGs). Os resultados são similares aos apresentados no estudo de 1997 (Iezzoni, 1997c), com as diferentes medidas a apresentarem valores distintos para as taxas de mortalidade observada, como traduzem os resultados do teste de Hosmer-Lemeshow e com as medidas baseadas nos resumos de alta a apresentarem um poder explicativo mais elevado.

Para este último aspecto avançam igualmente com uma explicação. O melhor valor explicativo das medidas baseadas nos resumos de alta pode depender da política de codificação, o que pode originar que, frequentemente, os doentes possam ser considerados mais “doentes” através desta modalidade, do que na realidade se observa quando se consultam os processos clínicos (Iezzoni, 1995).

Ainda a respeito da utilização de modelos de ajustamento pelo risco baseados nos dados administrativos, Iezzoni e outros (1992d) referem a importância da codificação. Neste artigo a autora defende que o preenchimento incompleto dos diagnósticos secundários nas bases de dados administrativas pode conduzir a enviesamentos na atribuição de risco, isto também porque muitas bases de dados somente permitem a consideração de 5 diagnósticos.

O estudo pretende avaliar o impacto de se ter mais diagnósticos nas bases de dados, apresentando os seguintes principais resultados e conclusões: alguns problemas de saúde, como a diabetes, enfarte do miocárdio antigo e angina, os quais seria de esperar que aumentassem o risco de morte, estão associados a riscos de morte baixos; o enviesamento ou diferentes práticas de codificação nas bases de dados administrativas podem contribuir para estes resultados, pelo que a intenção de se melhorar o registo dos diagnósticos secundários baseada somente no aumento dos campos preenchidos pode ser insuficiente.

Tendo em atenção estes aspectos Iezzoni (1997b) refere alguns requisitos para os Resumos Informatizados de Alta no futuro, nomeadamente os respeitantes aos elementos/variáveis que devem estar presentes, às metodologias e procedimentos de codificação, à necessidade de organizações independentes avaliarem todo este processo e ainda sobre o impacto provisional da alteração das modalidades de pagamento dos hospitais.

Assim, a escolha de um modelo de ajustamento pelo risco deve ter em conta os aspectos associados (Iezzoni, 1990 e 1997e e Griffith, Alexander e Jelinek, 2002):

- À significância clínica do sistema de classificação de doentes;
- À sua adaptabilidade às características de cada hospital;
- À credibilidade por parte dos médicos em relação ao seu desempenho para avaliar a qualidade dos cuidados prestados;
- À acessibilidade de dados em função do sistema de informação do hospital;
- À fiabilidade do sistema;
- Às exigências em termos da utilização de recursos por parte do hospital;
- Aos custos de exploração do sistema;
- À motivação e ao momento de escolha do sistema de classificação de doentes.

Neste sentido, interessa agora apresentar e discutir quais as possibilidades que neste momento se colocam em Portugal para a utilização e aplicação de sistemas de ajustamento pelo risco.

Existe um intenso debate em termos internacionais sobre a escolha dos melhores sistemas de ajustamento de risco, sendo de evidenciar os seguintes aspectos:

- Quais os melhores modelos? Os baseados em dados administrativos? Ou, os baseados em modelos clínicos?
- Deve escolher-se somente um modelo de ajustamento de risco para todas as finalidades? Ou deve escolher-se o melhor modelo para a respectiva finalidade?

Em relação ao primeiro aspecto, a dicotomia entre modelos administrativos e modelos clínicos, é conveniente apresentar previamente a sua caracterização. Modelos administrativos são, em termos internacionais, aqueles que utilizam os dados constantes nos resumos de alta. Na generalidade este tipo de suporte de dados contém elementos respeitantes à identificação do doente, à natureza da admissão, às transferências internas, ao destino após a alta, ao diagnóstico de admissão, à morfologia tumoral, ao peso à nascença, aos diagnósticos (até 20 e com indicação do diagnóstico principal), às causas externas de Lesão/Efeito Adverso (até 20), aos procedimentos (até 20), à data da primeira intervenção cirúrgica, ao número de dias em Unidades de Cuidados Intensivos e à identificação do médico responsável pelo tratamento e do médico codificador.

Os modelos clínicos são aqueles que incluem os dados dos modelos administrativos, acrescidos de alguns elementos constantes dos processos

clínicos, essencialmente aqueles que permitem caracterizar a história e o exame objectivo do doente, os resultados dos meios complementares de diagnóstico, tudo isto para se estabelecer um diagnóstico, estimar um prognóstico e prescrever o tratamento adequado.

Como foi referido o balanço que deve presidir a esta discussão deve ter em conta aspectos relacionados com a precisão e a validade do sistema, com os custos de implementação e de exploração destes sistemas e com o seu nível de operacionalização.

Enquanto que os aspectos associados à precisão e à validade dos sistemas de ajustamento pelo risco apontam para uma ligeira supremacia dos modelos clínicos, as questões relacionadas com os custos do sistema e a sua operacionalização referem que os modelos administrativos devem ser privilegiados.

Para Portugal, independentemente do valor intrínseco da precisão e da validade dos modelos é praticamente impossível aplicar, a curto prazo, os modelos clínicos pelo que se sugere a utilização exclusiva de modelos administrativos.

Esta opção não deve impedir que a médio prazo se altere o sistema de informação dos hospitais, de forma a considerar a eventualidade de se utilizarem os modelos clínicos.

No que se refere ao segundo aspecto, escolha de um ou mais modelos para ajustamento pelo risco em função da finalidade a que este se destina, a resposta pode ser mais flexível.

Assim, atendendo somente a este princípio deveria recomendar-se a utilização de diferentes modelos de ajustamento pelo risco, em função da finalidade a que o sistema se destina. O principal obstáculo consistirá então nos respectivos custos de implementação e de exploração dos referidos sistemas.

A este respeito a situação é mais pacífica. De facto, tendo em conta que se vão utilizar modelos administrativos, para os quais já existe actualmente colheita de dados e rotinas (que já envolvem portanto despesas), a utilização de mais do que um modelo representa apenas custos marginais insignificantes, pouco mais afinal do que aqueles decorrentes da aquisição de licenças para a utilização de programas informáticos.

Ou, ainda numa perspectiva mais positiva, a utilização de pelo menos dois sistemas permite rentabilizar o investimento feito na recolha de dados, visto que permite a introdução de metodologias mais correctas e amplas para se avaliar o desempenho dos hospitais.

Neste sentido, é importante identificar e caracterizar os principais sistemas de ajustamento pelo risco existentes, para finalmente se formularem recomendações sobre quais devem ser aplicados em Portugal (ver Quadro II).

Quadro II
 Principais Sistemas de Ajustamento pelo Risco

Sistema de Severidade	Dados Necessários	Definição de Severidade	Forma de Classificação
MedisGroups (versão original)	Dados Administrativos e Processo Clínico	Instabilidade Clínica indicada pela mortalidade	“Score” de Admissão, entre 0 e 4
MedisGroups (versão empírica)	Dados Administrativos e Processo Clínico	Mortalidade Hospitalar	Probabilidade de morte, entre 0 e 1
APACHE II (“score” fisiológico)	Dados Administrativos e Processo Clínico	Mortalidade Hospitalar para doentes das UCIs	“Score” entre 0 e 60
APACHE III (“score” fisiológico)	Dados Administrativos e Processo Clínico	Mortalidade Hospitalar para doentes das UCIs	“Score” entre 0 e 252
Computerized Severity of Illness	Dados Administrativos e Processo Clínico	Dificuldades no tratamento encontradas pelos médicos	“Score” entre 1 e 4 por doente e por cada código da CID-9-MC
Disease Staging (versão original)	Dados Administrativos	Estadio da doença baseado no risco de morte ou de incapacidade funcional	Três estadios, 1, 2 e 3, com sub-estadios dentro de cada estadio
Disease Staging (probabilidade de morte)	Dados Administrativos	Probabilidade de Morte Hospitalar	Probabilidade de morte, entre 0 e 1
Patient Management Categories	Dados Administrativos	Morbilidade e Mortalidade Hospitalar	“Score” entre 1 e 7
Índice de Comorbilidade (Índice de Charlson)	Dados Administrativos	Risco de morte dentro de 1 ano após a hospitalização	Valor único derivado de uma escala aditiva, representando o nº e a severidade das comorbilidades
Refined Diagnosis Related Groups	Dados Administrativos	Total da Facturação Hospitalar	4 escalões de severidade, dentro de cada DRG adjacente
All Patient Refined Diagnosis Related Groups	Dados Administrativos	Total da Duração de Internamento. Total da Facturação Hospitalar	3 classes de severidade dentro dos DRGs médicos; as mortes precoces são consideradas no grupo de severidade mais baixo

Fonte: Iezzoni, 1997b

Em Portugal, enquanto não for alterada a folha de admissão e alta e concomitantemente o resumo informatizado de alta, forma de recolha de dados que configura o que internacionalmente é conhecido por “dados administrativos”, torna-se impossível utilizar de forma sistemática a utilização de sistemas de severidade que necessitem de dados clínicos ou seja elementos referentes a

sintomas, sinais vitais e de disfunção fisiológica, os quais somente podem ser obtidos com consulta retrospectiva e exaustiva aos processos clínicos.

Por esta razão os sistemas como os MedisGroups, APACHE e Computerized Severity of Illness não podem ser utilizados.

Dos restantes, embora por razões diferentes, não devem ser utilizados os sistemas Patient Management Categories, e o Índice de Comorbilidade.

Os Patient Management Categories porque atendendo às suas características específicas, onde avulta a criação de um valor escalar que traduz o consumo de recursos necessários para cada categoria de doença, necessitam de uma adaptação específica ao estado da arte e à cultura portuguesa para o tratamento dos doentes.

O Índice de Comorbilidade porque considera a mortalidade após um ano de contacto com os hospitais, o que para além de suscitar alguns problemas técnicos (essencialmente a causalidade entre esta mortalidade e o tratamento hospitalar), é praticamente impossível de utilizar em Portugal, visto que as estatísticas de rotina não disponibilizam estes indicadores.

Assim, restam os sistemas de classificação Disease Staging e as versões Severity (SDRGs), Refined (RDRGs), All Patient Refined (APRDRGs) e International Refined dos Diagnosis Related Groups (IRDRGs).

Tendo em atenção as características destes sistemas de classificação de doentes, os quais para além de recorrerem a metodologias distintas para a identificação da severidade e de apresentarem produtos e resultados igualmente diferentes, apresentam, no entanto uma característica comum – utilizam os mesmos dados para identificarem produtos e estabelecerem medidas de ajustamento pelo risco, será interessante, no futuro considerar a utilização conjunta de uma das versões dos DRGs (IRDRGs) e os Disease Staging para se avaliar o desempenho dos hospitais portugueses.

No presente estudo, apesar de os aspectos referidos anteriormente, será utilizado unicamente um sistema de classificação de doentes para o ajustamento pelo risco – Disease Staging.

No entanto, embora não seja utilizado para se proceder ao ajustamento pelo risco, será igualmente utilizada a versão em vigor em Portugal dos DRGs, pelo que de seguida se irá fazer uma caracterização genérica destes dois sistemas de classificação de doentes.

Os Diagnosis Related Groups começaram a ser concebidos nos finais da década de sessenta na Universidade de Yale, com a principal finalidade de associar o

tipo de doentes tratados (“casemix”) com os respectivos custos (Fetter et al, 1980).

O desenvolvimento deste sistema de classificação de doentes obedeceu a uma série de requisitos, impostos pelos seus autores (Fetter et al, 1980 e Costa, 1994):

- O sistema de classificação teria de ser clinicamente interpretável, com grupos de diagnóstico homogêneos. Ou seja, os médicos deveriam estar aptos a relacionar os doentes de cada grupo com um determinado padrão de tratamento;
- A classificação deveria ser obtida mediante o recurso a informação generalizada e disponível nos hospitais;
- Os grupos constituídos pelo sistema de classificação deveriam ter um número finito, preferencialmente na ordem das centenas e serem exaustivos e mutuamente exclusivos;
- Cada grupo deveria incluir doentes com um consumo previsível de recursos similar;
- A definição de grupos deveria ser comparável entre os diversos sistemas de codificação.

Na primeira versão baseada na Classificação Internacional das Doenças – 8ª Revisão (CID-8) foram utilizadas seis variáveis – diagnóstico principal, diagnóstico secundário, procedimento cirúrgico principal, procedimento cirúrgico secundário, idade e presença ou ausência de cuidados psiquiátricos (Casas, 1991 e Rodrigues, 2002).

Em 1979, após a introdução da Classificação Internacional das Doenças, 9ª Revisão, Modificações Clínicas (CID-9-MC), introduziram algumas alterações, nomeadamente na redução das Grandes Categorias Diagnósticas (GCDs), na divisão entre casos médicos e cirúrgicos, na alteração dos critérios para a consideração da idade e dos diagnósticos secundários e na consideração das complicações (Casas, 1991 e Rodrigues, 2002).

Em 1983, os DRGs começaram a ser utilizados como sistema de pagamento prospectivo dos hospitais americanos, pelo que atendendo a esta finalidade davam especial atenção à intensidade na utilização de recursos (Averill et al, 1998).

Contudo, as alterações registadas nos sistemas de saúde, nomeadamente nos EUA, aumentaram o âmbito de aplicação dos sistemas de classificação de doentes e dos DRGs, para dimensões como (Averill et al, 1998):

- Comparar os hospitais, tanto em termos de recursos como de resultados;
- Avaliar as diferenças entre taxas de mortalidade;

- Facilitar a implementação de critérios de decisão clínica;
- Servir de suporte à gestão e ao planeamento.

Estes aspectos conduziram ao aparecimento de diversas modalidades de DRGs, dentro das quais se devem referir os Medicare DRGs, os Refined DRGs (RDRGs), os All Patient DRGs (APDRGs), os SDRGs, os APRDRGs e os IRDRGs (Averill et al, 1998 e Mullin, Averill e Boucher, 2002).

A descrição genérica dos DRGs, bem como as suas principais vantagens e inconvenientes encontram-se bem identificados noutros artigos (Fetter et al, 1980; Hornbrook, 1982; Aronow, 1988; Urbano e Bentes, 1990; Costa e Nogueira, 1994; Costa, 1994 e Bentes et al, 1996), pelo que neste momento se fará somente uma caracterização genérica dos DRGs, embora dando especial importância aos sistemas desenvolvidos pela Health Care Financing Administration (HCFA) e aos SDRGs, APRDRGs e IRDRGs.

Os elementos constantes no Quadro III permitem visualizar uma comparação rápida dos diferentes sistemas de DRGs (Averill et al, 1998).

Quadro III
 Comparação dos Diferentes Sistemas de DRGs

	Medicare DRGs (Versão12)	SDRGs	RDRGs (Versão 10)	APDRGs (Versão 12)	APRDRGs (Versão 12)
Nº de categorias de DRG	338	316	367	ND	384
Nº DRGs	492	652	1170	641	1530
GCD (a) para Politraumatizados	Limitado	Limitado	Limitado	Completo	Completo
GCD para VIH (b)	Limitado	Limitado	Limitado	Completo	Completo
Peso à nascença	Não	Não	Limitado	Completo	Completo
Indicador Pediátrico	Não	Não	Não	Limitado	Completo
Grandes Complicações e Comorbilidades	Não	Sim	Sim	Sim	Sim
Morte utilizada na definição	Sim	Sim	Sim	Sim	Não
Duração de Internamento utilizada na definição	Não	Não	Sim	Somente p/ Recém Nascidos	Não
Reavaliação da Lista de Complicações e Comorbilidades	Não	Grande	Não	Limitada	Completa
Reconhecimento de Complicações e Comorbilidades Múltiplas	Não	Não	Não	Não	Sim
Estrutura para sub-grupos de Complicações e Comorbilidades	Variável	Variável	Relativamente Uniforme	Variável	Uniforme
Número de sub-grupos de Complicações e Comorbilidades	2	3	3 Médicos; 4 Cirúrgicos	3	4
Sub-grupos de Risco de Morte	Não	Não	Não	Não	Sim
Base de Dados utilizada	-----	Medicare	Medicare	Medicare	APDRGs, excepto para recém nascidos

Fonte: Averill et al, 1998

(a) – GCD – Grandes Categorias Diagnósticas

(b) – VIH – Vírus da Imunodeficiência Humana

As principais diferenças encontradas entre os diversos sistemas de DRGs verificam-se no número de categorias de cada sistema, na consideração mais ou menos completa de Grandes Categorias Diagnósticas (GCD) para Politraumatizados e para doentes infectados com o Vírus da Imunodeficiência Humana (VIH), para a forma de reconhecimento e para os grupos destinados às Complicações e às Comorbilidades e à introdução de sub-grupos de Risco de Morte.

Os SDRGs foram iniciados com uma reavaliação efectuada pela Health Care Financing Administration (HCFA) em 1993, para a qual foi utilizada a Lista de Grandes Complicações e Comorbilidades concebida para os APDRGs, com excepção das grávidas, recém nascidos e doentes pediátricos (Averill et al, 1998). O resultado final foi a definição de 111 grupos de diagnóstico sem complicações e comorbilidades, 220 grupos de diagnósticos nos quais a presença de complicações e comorbilidades não foi considerada suficiente para a criação de sub-grupos de complicações e comorbilidades e 395 grupos de diagnóstico nos quais as complicações e as comorbilidades foram consideradas relevantes. A lista de base dos DRGs foi então subdividida em nenhum, dois ou três sub-grupos, sendo o resultado final de 84 DRGs isolados, 124 com dois sub-grupos e 85 com três sub-grupos. Assim, adicionando estes aos DRGs não considerados inicialmente o total de DRGs foi de 652. A lógica do sistema considera que a um doente é sempre atribuído o sub-grupo correspondente ao mais alto nível entre os diagnósticos secundários. Este sistema não é actualizado desde 1994 (Averill et al, 1998).

Os APRDRGs constituem uma melhoria do sistema APDRGs, com a atribuição de 4 sub-grupos suplementares. Para tal todos os elementos respeitantes à idade e grandes complicações e comorbilidades considerados nos APDRGs foram substituídos por dois conjuntos de sub-grupos. Um primeiro respeitante à gravidade da doença e o segundo em relação ao risco de morte. Para cada um destes conjuntos são considerados 4 escalões que representam valores reduzidos, moderados, maiores e extremos de gravidade ou risco de morte. A classificação dos doentes nestes escalões é efectuada em função dos diagnósticos secundários, bem como da sua interacção, idade, diagnóstico principal e a existência de alguns procedimentos não cirúrgicos. Esta classificação é feita separadamente para a gravidade e para o risco de morte. Este processo conduziu à criação de 1530 grupos de diagnóstico, na qual a interacção entre diagnósticos secundários e a existência de diversas doenças, envolvendo ainda diversos sistemas orgânicos, constitui o elemento essencial na identificação de doentes com problemas mais graves (Averill et al, 1998).

Os IRDRGs, um dos sistemas mais recentes dos DRGs, consistem numa adaptação e melhoria dos APDRGs e dos APRDRGs, tendo sido apresentadas como principais razões para o seu aparecimento, a necessidade de se criar um sistema de DRGs único para todos os países, a capacidade deste sistema para se fazerem comparações internacionais e a necessidade de o sistema de

classificação de doentes se adaptar a diferentes modelos de codificação (Mullin, Averill e Boucher, 2002).

Os IRDRGs, formas construídos recorrendo a dados dos EUA, da Bélgica, da Itália e da Espanha. Utilizando a mesma lógica dos APDRGs e dos APRDRGs, são considerados três níveis de severidade, para a grande maioria dos DRG. Esta atribuição está baseada na presença e na gravidade das complicações e das comorbilidades, correspondendo cada um dos grupos, respectivamente a ausência de complicações e comorbilidades, à sua presença, ou ainda à existência de grandes complicações e comorbilidades (Mullin, Averill e Boucher, 2002).

Segundo estes autores, cada nível de severidade traduz o consumo de recursos de cada doente, sendo para tal determinante a presença e os efeitos dos diagnósticos secundários.

Assim, os IRDRGs consistem em 319 DRGs de base, dos quais 301 apresentam três níveis de severidade, 9 DRGs sem qualquer nível de severidade, 3 DRGs cirúrgicos não relacionados com o diagnóstico principal, igualmente com três níveis de severidade, a que se adicionam dois DRGs residuais, pelo que no total existem 923 DRGs (Mullin, Averill e Boucher, 2002).

O Disease Staging é um sistema de classificação de doentes que pretende responder às questões colocadas pelos médicos quando analisam a situação de saúde de um doente, designadamente, onde? porquê? e qual a gravidade da situação? (Gonnella, Louis e Gozum, 1999). Segundo estes autores, a localização identifica o sistema orgânico afectado, o porquê refere-se à etiologia do problema e a gravidade dá indicações sobre as alterações fisiológicas decorrentes das complicações da doença.

Neste sentido referem ainda que os médicos utilizam a informação retirada da história do doente, do exame objectivo do doente e dos resultados dos meios complementares de diagnóstico para responderem a estas três questões e para estabelecerem o diagnóstico, estimarem um prognóstico e prescreverem o tratamento adequado (Gonnella, Louis e Gozum, 1999).

Nesta linha de pensamento o Disease Staging é um sistema de classificação de doentes que utiliza os elementos constantes no resumo de alta (dados administrativos), designadamente o diagnóstico principal e os diagnósticos secundários, para produzir grupos de doentes que requerem tratamentos idênticos e com idênticos resultados esperados (Garg et al, 1978; Gonnella, Hornbrook e Louis, 1984 e Gonnella, Louis e Gozum, 1999).

Atendendo a estes princípios espera-se que constitua um sistema que propicie metodologias para a avaliação da qualidade dos cuidados prestados, incluindo a

avaliação dos resultados em saúde, a análise da eficiência e da eficácia/efectividade dos tratamentos e para o “benchmarking” e a avaliação do desempenho hospitalar.

Para tal, a versão clínica dos Disease Staging inclui 4 estadios (Gonnella, Louis e Gozum, 1999 e MEDSTAT, 2001):

- Estadio 1 – Doença sem complicações;
- Estadio 2 – Doença com complicações locais;
- Estadio 3 – Doença que compreende diferentes localizações, ou complicações sistémicas;
- Estadio 4 – Morte.

Para cada estadio existem sub-estadios, tendo essencialmente em vista a obtenção de uma medida mais precisa. Como principal particularidade deve-se ter em atenção que para algumas doenças é igualmente considerado um estadio 0, o qual pretende identificar riscos que podem conduzir ao aparecimento da doença, como por exemplo no caso de história/antecedentes pessoais de carcinoma (Gonnella, Louis e Gozum, 1999).

É ainda de referir que este sistema de classificação utiliza uma escala ordinal e que os mesmos estadios entre doenças não disponibilizam a mesma informação, nem estabelecem o mesmo risco de um resultado indesejável (Hornbrook, 1982 e Thomas, Ashcraft e Zimmerman, 1986).

Por outro lado, a criação de sub-estadios não é uniforme, nem em termos de categorias nem na utilização de toda a escala. Enquanto que, por exemplo, o Hipertiroidismo (Etiologia: Metabólica – ND15) inclui 4 estadios, com 10 sub-estadios, a Diabetes mellitus (Etiologia: Metabólica – ND12) inclui igualmente 4 estadios, mas com 17 sub-estadios. Em relação à utilização total da escala, podem citar-se os exemplos das Cataratas, para as quais os estadios 3 e 4 não são considerados, ou da Meningite bacteriana, em que o estadiamento começa no nível 2 (Gonnella, Louis e Gozum, 1999).

Após todas as doenças estarem identificadas e faseadas desenvolveu-se o Coded Staging Criteria, o qual corresponde genericamente, à atribuição para cada doença e respectivo estadio e sub-estadio de códigos de doença de acordo com a Classificação Internacional das Doenças – 9ª Revisão – Modificações Clínicas (CID-9-MC) (Gonnella, Louis e Gozum, 1999 e MEDSTAT, 2001) (ver no Anexo I alguns exemplos).

Após esta fase foi definida uma Categoria Principal de Doença (PDXCAT), bem como um valor principal de estadio e sub-estadio (Gonnella, Louis e Gozum, 1999 e MEDSTAT, 2001).

Como somente é atribuída uma PDXCAT para cada admissão, sendo esta baseada no diagnóstico principal definido no resumo de alta foi definida a seguinte metodologia para se identificarem os estadios dos diagnósticos principais e secundários (Gonnella, Louis e Gozum, 1999 e MEDSTAT, 2001):

- Quando o diagnóstico secundário é considerado como uma complicação do diagnóstico principal, esta pode ser considerada para alterar o estadio do diagnóstico principal (por exemplo, quando a Diabetes mellitus é o diagnóstico principal e se está na presença de retinopatia e de neuropatia, como diagnósticos secundários, então estes últimos são considerados complicações do diagnóstico principal e o “software” identificará uma PDXCAT e um estadio específico);
- Quando o diagnóstico secundário não está associado com o diagnóstico principal, é então definida uma Categoria Secundária de Doença (DXCAT), bem como o respectivo estadio deste diagnóstico secundário;
- Quando os diagnósticos secundários são sintomas, então nenhuma DXCAT é considerada.

Para qualquer destas versões o Disease Staging apresenta uma escala ordinal para cada doença, sendo ainda de evidenciar que os escalões não são equivalentes entre as diversas doenças (Gonnella, Louis e Gozum, 1999).

Por estas razões a MEDSTAT desenvolveu um “software” específico para os Disease Staging, o qual após a codificação dos Disease Staging, foi desenvolvido para se atribuírem diagnósticos principais e estadios para cada doente, em função dos dados incluídos nos resumos de alta de cada hospital (MEDSTAT, 2001).

Esta versão automatizada (MEDSTAT, 2001), permite identificar os estadios e sub-estadios referidos na versão clínica e original, apresenta previsões, numa escala de intervalos, para a duração de internamento, para os custos, para a mortalidade, para as complicações e para as readmissões.

O “software” do Disease Staging, atentos os problemas da escala ordinal e de comparação entre doenças referidos para a versão clínica, produz, na generalidade três tipos de indicadores: Indicadores de Recursos, Indicadores de Demora Média e Indicadores de Mortalidade (MEDSTAT, 2001). Estes indicadores para além de obviarem aos problemas de comparação presentes na versão clínica dos Disease Staging, permitiram o aparecimento de metodologias para se avaliar a efectividade e a eficiência das organizações de saúde.

Assim, a utilização combinada das duas versões permite, segundo os seus autores (Gonnella, Louis e Gozum, 1999), a aplicação do Disease Staging nas seguintes áreas:

- Apropriação da hospitalização;
- Análise da utilização de recursos;
- Financiamento;
- Avaliação da qualidade dos cuidados prestados;
- Ensaios clínicos;
- Planeamento de recursos;
- Certificação dos recursos;
- Educação médica.

Finalmente deve ser referido um aspecto ainda não debatido neste artigo: as características dos hospitais devem ou não ser consideradas em modelos para ajustamento pelo risco?

A grande maioria dos autores defende a sua exclusão (Fetter et al, 1980; Young, Swinkola e Zorn, 1982; Knaus et al, 1986; Horn, 1988; Gonnella et al, 1990; Le Gall, Lemeshow e Saulnier, 1993; Lemeshow et al, 1993 e Brewster et al, 1995).

Os principais argumentos para esta exclusão encontram-se bem sintetizados no artigo de DesHarnais e outros (1997). Neste artigo defende-se que as características dos hospitais (tipo de propriedade, a dimensão, a função de ensino e a existência de programas de formação profissional, bem como a localização, urbana ou rural e a taxa de ocupação dos hospitais) não devem ser utilizados no ajustamento pelo risco, essencialmente porque estas dimensões não qualificam o risco clínico de cada doente. Antes pelo contrário estas características representam o risco institucional associado ao facto de um doente ser admitido num determinado hospital. Enfatizam ainda a importância de se excluirmos as características dos hospitais, simplesmente pelo facto de poderem ser atribuídos níveis de risco diferentes a doentes com situações clínicas idênticas, em função do tipo de hospital onde este recebeu tratamento.

2.3.3.6. Propriedades dos Modelos de Ajustamento pelo Risco

Na sua grande maioria os modelos de ajustamento pelo risco têm recorrido a técnicas de regressão logística multivariada para se estabelecerem previsões sobre o resultado em análise (Iezzoni, 1997b).

Em geral os modelos de regressão constituem um método importante para a análise multivariada de dados, visto que permitem traduzir a relação existente entre uma variável dependente ou resultado e várias variáveis explicativas.

Neste particular, os modelos de regressão logística representam um método comum para a análise de relações em que a variável dependente não é contínua (Hosmer e Lemeshow, 1989).

Esta é a situação mais frequente quando se pretende analisar a mortalidade hospitalar, em que a variável dependente é dicotómica, assumindo os valores 0 (alta vivo) e 1 (alta falecido) em função de um conjunto de variáveis independentes, frequentemente designadas por covariáveis (Hosmer e Lemeshow, 1989; Ruttiman, 1994).

As características técnicas, bem como os procedimentos necessários para se realizarem regressões logísticas encontram-se bem documentados em diversos estudos e publicações (Greenland, 1984; McNeil e Hanley, 1984; Hosmer e Lemeshow, 1989; Hosmer, Jovanovic e Lemeshow, 1989; Pryor e Lee, 1991; Ruttimann, 1994 e Smith et al, 1996) pelo que somente se evidenciarão os aspectos relacionados com o ajustamento do modelo.

Segundo Ruttiman (1994) a regressão logística apresenta as seguintes vantagens: (1) não é necessário introduzir muitas restrições sobre a distribuição das covariáveis e (2) o modelo logístico disponibiliza estimativas directas sobre a probabilidade de ocorrência do “outcome”.

Neste sentido, após a estimação dos coeficientes de regressão logística para as covariáveis com significância estatística deve ter-se em atenção o seguinte (Hosmer e Lemeshow, 1989):

- Pode estimar-se o respectivo “odds-ratio” para cada uma das covariáveis;
- Deve proceder-se à avaliação do ajustamento do modelo, o que deverá somente ocorrer quando existir a convicção que o modelo final contém todas as covariáveis, interações incluídas, pertinentes para a previsão da variável dependente.

O ajustamento do modelo deve ser avaliado em função de dois aspectos:

1. A calibração;
2. A discriminação.

Na calibração analisam-se os desvios entre a mortalidade observada e a mortalidade prevista pelo modelo, por outras palavras comparam-se os valores médios entre a mortalidade observada e os valores médios para a probabilidade prevista pelo modelo (Ash e Shwartz, 1997).

Neste sentido, a afirmação de que um modelo está bem calibrado baseia-se no pressuposto de que todos os parâmetros incluídos no modelo derivam do respectivo ajustamento do modelo aos dados. Quando se utiliza o método dos mínimos quadrados é natural que o erro de calibração seja mínimo. Contudo, quando se utiliza outra abordagem, como por exemplo, a da máxima verosimilhança é natural que existam diferenças entre os valores médios observados e os valores médios previstos (Ash e Shwartz, 1997).

A discriminação analisa o facto de o sistema prever taxas de mortalidade mais elevadas para os doentes que efectivamente morrem, comparativamente com os doentes que efectivamente não morrem (Ash e Shwartz, 1997).

Existe um intenso debate sobre qual a característica mais importante para determinar o nível de ajustamento de um modelo.

Enquanto, Lemeshow e Hosmer (1982) defendem que se um modelo não está bem calibrado é inútil avaliar a sua discriminação Harrel e outros (1984) referem que a discriminação é a característica mais importante de um modelo, visto que a calibração pode ser atingida posteriormente com alguns ajustamentos subjectivos, derivados da integração no modelo das opiniões dos peritos.

Isto porque, se um modelo tem boa discriminação a calibração pode ser alcançada sem prejuízo daquele atributo, enquanto que a ausência de poder de discriminação nunca pode ser corrigida com melhorias na sua calibração (Harrel et al, 1984).

Por outro lado, segundo Ash e Shwartz (1997), quando se pretende diferenciar somente o risco de morte, ou seja, a distinção entre mortos e sobreviventes, a calibração não é um aspecto importante. O mesmo já não deve ser afirmado quando se pretendem comparar os valores da mortalidade prevista com a observada, essencialmente para efeitos de garantia de qualidade com o estabelecimento de “casos-sentinela”, em que a calibração assume um carácter muito importante.

Calibração

Existem diversas alternativas para se apurar o ajustamento dos modelos de regressão logística (Hosmer e Lemeshow, 1989; Wagner, Knaus e Draper, 1983; Ruttimann, 1994 e Ash e Shwartz, 1997):

- O Qui-Quadrado de Pearson;
- A “Deviance”;
- O teste de Hosmer-Lemeshow;
- O coeficiente de determinação (R^2).

Neste estudo, à semelhança de vários estudos internacionais, será somente analisado o teste de Hosmer-Lemeshow, baseando-se este teste no agrupamento das probabilidades de morte estimadas (Lemeshow e Hosmer, 1982).

Existem dois tipos de problemas quando se utiliza esta estatística (Ash e Shwartz, 1997):

- Quando os sub-grupos formados (correspondentes aos decis das probabilidades de morte estimadas) apresentam probabilidades de morte com grande dispersão, o valor da estatística do teste de Hosmer-Lemeshow é subavaliado, conduzindo a uma aceitação inapropriada do modelo;
- No teste de Hosmer-Lemeshow, à semelhança do teste do qui-quadrado, a aceitação do modelo depende fortemente do número de observações. Quando a dimensão da amostra é grande, mesmo que se observem discrepâncias pequenas entre a mortalidade observada e prevista, é provável que o modelo seja rejeitado, não acontecendo o mesmo na situação inversa, ou seja quando a dimensão da amostra é pequena, pode acontecer que o modelo seja aceite, inclusivamente mesmo quando se está na presença de discrepâncias entre as probabilidades de morte observada e prevista.

Apesar dos potenciais problemas referidos, o teste de Hosmer-Lemeshow tem sido uma das estatísticas mais utilizadas para se apurar a calibração do modelo.

Discriminação

Conforme foi referido a discriminação detecta se o sistema prevê taxas de mortalidade mais elevadas para os doentes que efectivamente morrem, comparativamente com os doentes que efectivamente não morrem (Ash e Shwartz, 1997).

Na generalidade existem duas formas para se averiguar o nível de discriminação de um modelo (Harrel et al, 1984; Knaus e Harrel, 1989; Harrel et al, 1990; Carneiro, 1994; Ruttimann, 1994; Wagner, Wagner et al, 1994 e Ash e Shwartz, 1997):

- as curvas ROC (“receiver operating characteristics curves”);
- A estatística “c”.

A curva ROC é obtida pela representação gráfica para todos os valores possíveis da probabilidade de morte prevista da sensibilidade versus 1-especificidade (proporção de falsos positivos).

Em que, atendendo aos valores do Quadro IV (Ruttimann, 1994 e Ash e Shwartz, 1997):

Quadro IV
 Comparação entre um Modelo com Resultados Dicotómicos
 e um Modelo com Previsões Dicotómicas

Previsões de risco do modelo	Resultados		
	Falecidos	Sobreviventes	Total
Falecidos	A	B	A+B
Sobreviventes	C	D	C+D
Total	A+C	B+D	A+B+C+D

- verdadeiros positivos = A
- falsos positivos = B
- verdadeiros negativos = D
- falsos negativos = C
- prevalência = $(A+C) / (A+B+C+D)$
- precisão (eficiência global) = $(A+D) / (A+B+C+D)$
- sensibilidade = $A / (A+C)$
- especificidade = $D / (B+D)$
- proporção de falsos negativos = $1 - \text{sensibilidade}$
- proporção de falsos positivos = $1 - \text{especificidade}$
- valor predictivo positivo = $A / (A+B)$
- valor predictivo negativo = $D / (C+D)$

Em que para um determinado “cut-off point” - “t” - quando as probabilidades de morte estimadas são superiores a “t” todos os doentes devem falecer e quando as probabilidades de morte são inferiores a “t” é esperado que todos os doentes sobrevivam (Ruttimann, 1994 e Ash e Shwartz, 1997).

Neste sentido, quando “t” = 0, todos os casos são positivos e o par (1,1) é gerado e quando “t” =1, todos os casos são negativos, gerando-se o par (0,0).

Assim, espera-se a formação de um gráfico da Curva ROC (ver Figura 3).

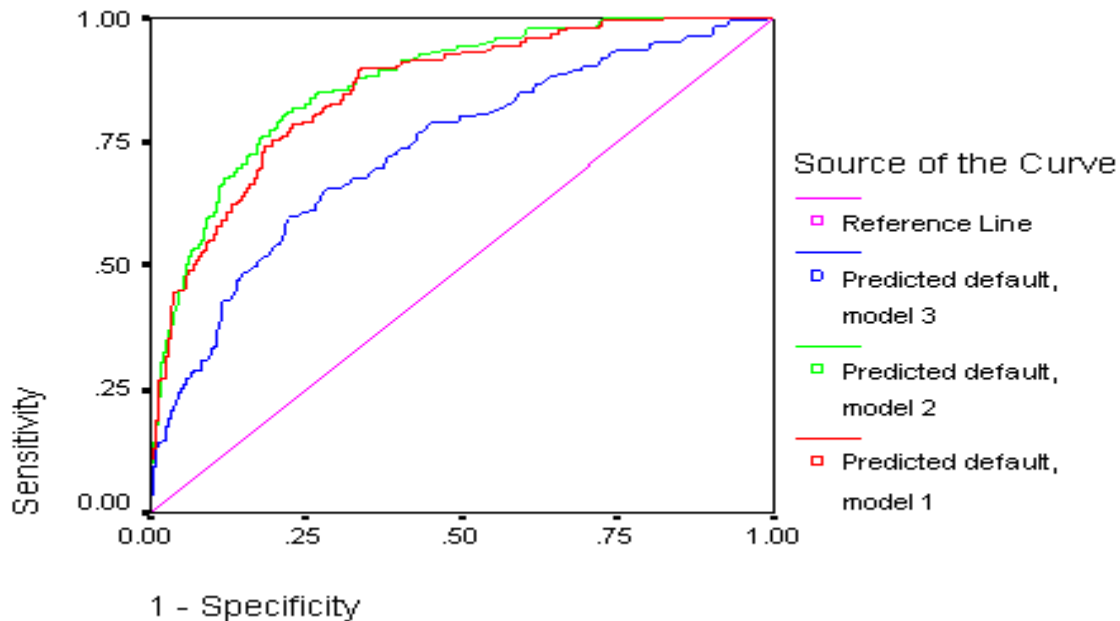
Os valores da diagonal entre os pares (0,0) e (1,1) exprimem que os mesmos são obtidos por acaso, ou seja, que o modelo não apresenta um nível de discriminação. Quanto mais a curva ROC se afastar da diagonal, maior será o poder discriminativo do modelo (como por exemplo é ilustrado na Figura 3).

Mas, embora as curvas ROC forneçam uma indicação visual sobre o poder discriminativo do modelo é conveniente disponibilizar uma informação com maior precisão. É principalmente por esta razão que se deve calcular a estatística “c”.

A estatística “c” é igual à área sob a curva ROC e assume valores entre 0.5 e 1, correspondendo 0.5 a um nível discriminativo mínimo e 1 a um poder de discriminação máximo (Ruttimann, 1994).

A interpretação do valor da estatística “c” é a seguinte (Ruttimann, 1994): um valor de 0.9, por exemplo, permite concluir que em 90% das situações, os doentes falecidos apresentam um índice de severidade superior ao dos doentes que sobreviveram, mas não significa que a previsão dos falecidos tem uma probabilidade de 0.9 ou que a previsão de morte está associada com os falecimentos observados em 90% das situações.

Figura 4
Curva ROC (exemplo demonstrativo)



Fonte: SPSS, Versão 11.5

2.3.3.7. Validação dos Modelos de Ajustamento pelo Risco

Existem diversos problemas quando se pretendem aplicar sistemas de ajustamento de risco a populações diferentes ou a sub-grupos distintos das populações de onde o sistema foi originalmente construído (Iezzoni, 1999).

Num estudo realizado por Justice, Covinsky e Berlin (1999) são referidos alguns aspectos conceptuais e metodológicos respeitantes à utilização de modelos de ajustamento pelo risco (ver Figura 5).

Atendendo a que existem problemas na aplicação indiscriminada de modelos a diferentes populações, os autores descrevem uma abordagem baseada na precisão dos modelos (calibração e discriminação) e na generalização dos modelos (reprodutibilidade e transportabilidade).

Como subsistem alguns problemas semânticos, os autores avançam com definições de precisão e de generalização (ver Figura 5) as quais são

consideradas como úteis para a validação de modelos de ajustamento pelo risco a populações distintas (Justice, Covinsky e Berlin, 1999).

Figura 5
 Aspectos Conceptuais e Metodológicos dos Modelos de Ajustamento pelo Risco

Termo	Definição ou Critério
Precisão	O nível segundo o qual os valores previstos estão em conformidade com os observados
Calibração	A probabilidade prevista nunca é demasiado alta, nem demasiado baixa
Discriminação	A ordenação dos riscos individuais está correcta
Generalização	Capacidade de o modelo de ajustamento pelo risco estabelecer previsões precisas numa população diferente
Reprodutibilidade	O sistema é preciso em doentes não incluídos no seu desenvolvimento, mas que são de uma população idêntica
Transportabilidade	O sistema é preciso em doentes de uma população diferente, embora relacionada com a original
Histórica	A precisão mantém-se em períodos de tempo diferentes
Geográfica	A precisão mantém-se quando o sistema é testado em localizações diferentes
Metodológica	A precisão mantém-se quando o sistema é testado em métodos de recolha de dados distintos
Espectro	A precisão mantém-se em populações que, em média, estão mais ou menos avançadas na progressão da doença ou com uma trajetória diferente de doença
“Follow-up”	A precisão mantém-se quando o sistema é testado em períodos curtos ou longos

Os autores (Justice, Covinsky e Berlin, 1999) propõem ainda uma hierarquia para a validação externa de sistemas de ajustamento pelo risco (ver Figura 6).

Figura 6
 Hierarquia para a Validação Externa de Sistemas de Ajustamento pelo Risco

Nível de Validação	Avaliação Cumulativa da Generalização
0. Interna	Reprodutibilidade
1. Prospectiva	Reprodutibilidade; Transportabilidade Histórica
2. Independente	Reprodutibilidade; Transportabilidades Histórica, Geográfica, Metodológica e de Espectro
3. Localização Múltipla	Reprodutibilidade; Transportabilidades Histórica, Geográfica, Metodológica e de Espectro
4. Independente e Múltipla	Reprodutibilidade; Transportabilidades Histórica, Geográfica, Metodológica e de Espectro
5. Independente e Múltipla para Tabelas de Sobrevida	Reprodutibilidade; Transportabilidades Histórica, Geográfica, Metodológica, de Espectro e de “Follow-up”

Num estudo realizado por Flanders e outros (1999) pretende-se avaliar a discriminação e a calibração do “Pneumonia Severity-of-Illness” (PSI), uma medida de prognóstico específica por doença, quando aplicado aos doentes com “Community-acquired pneumonia” (CAP).

Os principais resultados referem que o PSI apresenta uma boa discriminação entre doentes com alto risco de morte e com aqueles que apresentam baixo risco de morte.

Contudo, a calibração do modelo foi fraca, tendo o PSI previsto 2.4 vezes mais mortes do que aquelas que ocorreram nos doentes com CAP.

Para tal são avançadas algumas explicações, dentro das quais avulta a questão de o modelo não apresentar boa calibração porque existem alguns factores que não estão a ser devidamente valorizados e medidos. Estes factores poderão ser as diferenças organizacionais e regionais nos critérios de admissão de doentes, as diferenças no tratamento e na disponibilidade tecnológica por parte dos hospitais e as diferenças nos próprios factores de prognóstico.

Neste sentido, os autores apresentam uma metodologia para a recalibração do modelo, a qual se baseou na regressão logística de acordo com o recomendado nos estudos de lezzoni (1993 e 1996b), referindo ainda que a recalibração é necessária sempre que se aplicam resultados a uma população diferente.

Finalmente deve ser referido um artigo de lezzoni (1999), no qual a autora baseando-se nos resultados de Flanders e outros (1999) e de Justice, Covinsky e Berlin (1999) refere que sempre que se transportam resultados de medidas estatísticas para outras populações, tanto em termos espaciais, como temporais, os analistas devem observar o princípio “Caveat emptor”. Ou seja, muito dificilmente o modelo aplicado a diferentes populações apresentará resultados precisos. A autora refere ainda que estes problemas de calibração podem ser solucionados com recurso a técnicas estatísticas como a regressão logística.

Para melhor se compreender a dimensão dos fenómenos e se definir uma metodologia para a sua solução, a autora sugere os seguintes princípios (lezzoni, 1999):

- Identificação precisa da versão dos sistemas de classificação de doentes que está a ser utilizada no estudo;
- Sempre que for conveniente, os autores devem utilizar os mecanismos mais adequados de diagnóstico do modelo. No caso dos sistemas de ajustamento pelo risco, deve sempre avaliar-se a discriminação e a calibração do modelo;
- Quando se realizarem processos de recalibração os autores devem referir detalhadamente todos os passos que foram utilizados na técnica.

Mesmo que se cumpram estas fases, a autora chama a atenção para o facto de que a recalibração é sempre um fenómeno único, visto que o novo modelo somente está ajustado aos dados e à população que serviu de base para o estudo.

Assim, pode concluir-se que embora exista uma enorme variedade de modelos para a avaliação do desempenho, as quais configuram diversas dimensões e perspectivas, existem igualmente alguns elementos comuns, designadamente a necessidade de combinar as características dos doentes, com diversos procedimentos estatísticos, tendo em vista o ajustamento pelo risco.

Por sua vez, o ajustamento pelo risco é considerado como um procedimento inevitável e incontornável de uma abordagem correcta para avaliar o desempenho hospitalar.

OBJECTIVOS

3. Objectivos

A avaliação da produção e do desempenho hospitalar têm constituído uma das maiores preocupações nas áreas de política de saúde e de administração em saúde.

Para tal contribuem a natureza específica do produto de saúde e naturalmente do produto hospitalar, bem como a existência de diferentes, se não conflitantes, perspectivas e funções de utilidade dos agentes que operam no mercado da saúde.

Por outro lado, existem algumas áreas de investigação e de aplicação que emergem sempre que se discute a temática hospitalar, designadamente o financiamento hospitalar, a qualidade e eficiência dos cuidados prestados e, eventualmente, a “saúde” financeira de todas as organizações de saúde.

Embora estas áreas por vezes justifiquem a existência de tensões e de conflitos entre os diversos agentes do sector da saúde, nomeadamente sobre a supremacia de determinadas abordagens, metodologias e procedimentos para se cumprirem os objectivos e finalidades da política, da regulação e da administração em saúde, um elemento parece consensual nesta panorama – a necessidade de se identificar, medir e comparar a produção hospitalar.

No entanto, a utilidade deste assunto é ainda potenciada quando se perspectiva que a identificação e medição da produção hospitalar, para além do valor intrínseco, é ainda o elemento comum para o financiamento e para a avaliação do desempenho dos hospitais.

É neste contexto que se desenvolvem abordagens mais funcionais, em que as preocupações dominantes respeitam à identificação de produtos que permitam uma aplicação financeira e administrativa, em oposição a outras mais conceptuais, cujas principais preocupações são movidas para aspectos mais técnicos relacionados com a identificação de produtos que permitam a sua aplicação no processo de diagnóstico e prognóstico da situação, bem como da definição dos modelos terapêuticos adequados que permitam a optimização da qualidade e dos ganhos em saúde.

A tradução prática destas questões conduziu ao desenvolvimento de sistemas de classificação de doentes mais preocupados com a complexidade dos casos tratados, cujo expoente máximo são os Diagnosis Related Groups (DRGs) ou, a outros que pretendem medir a gravidade dos doentes, como por exemplo, os MedisGroups, o Disease Staging e algumas versões mais recentes dos DRGs, nomeadamente os All Patient Refined (APRDRGs) e os International Refined ICD-9-CM (IRDRGs).

Face ao exposto e à realidade existente em Portugal, na qual somente está a ser aplicado um sistema de classificação de doentes, os DRGs, predominantemente centrado na complexidade dos casos e cuja principal finalidade é o financiamento dos hospitais, torna-se igualmente importante investigar a possibilidade de introdução de um sistema de classificação que identifique e meça a gravidade dos doentes.

Tendo como cenário de fundo as limitações do sistema de informação hospitalar português, em que predominam o que internacionalmente é designado por modelos administrativos, neste estudo será utilizado o Disease Staging para responder às insuficiências existentes na identificação e medição da gravidade.

Deve igualmente referir-se que para se responder às necessidades de medição da gravidade, poderiam ter sido utilizadas as versões APRDRGs ou IRDRGs. No entanto, esta não foi opção prioritária, visto que se privilegiou a comparação de sistemas que têm abordagens completamente distintas – a funcional (DRGs) e a conceptual – Disease Staging.

Contudo, atendendo à importância do tema é natural afirmar que no futuro deverão ser realizados outros estudos com a utilização de uma das versões dos DRGs que contempla a gravidade, para se compararem esses resultados com os disponibilizados pelo Disease Staging e, desta forma, se aumentar o conhecimento da produção e do desempenho dos hospitais em Portugal.

Assim, o presente estudo vai analisar os aspectos relacionados com a monitorização e avaliação da produção e do desempenho hospitalar, mediante a optimização do sistema de informação hospitalar existente em Portugal, no que respeita aos resumos de alta do internamento.

O estudo apresenta as seguintes principais finalidades:

- Contribuir para a melhoria dos métodos para a identificação e caracterização da produção e do desempenho hospitalar;
- Contribuir para uma maior racionalização no financiamento e na regulação dos hospitais, designadamente através da utilização de indicadores que caracterizem a procura deste tipo de cuidados;
- Contribuir igualmente para a melhoria da gestão dos hospitais, nomeadamente através da utilização de novos mecanismos e de instrumentos que monitorizem a sua actividade e funcionamento;
- Contribuir para uma melhor prestação de cuidados às populações, quantitativa e qualitativa, tendo em vista a obtenção de ganhos em saúde.

Neste sentido, os objectivos do estudo são os seguintes:

- Utilizar a gravidade para avaliar a produção e o desempenho hospitalar;
- Utilizar o conceito de ajustamento pelo risco para avaliar o desempenho dos hospitais;
- Definir e operacionalizar algumas dimensões para acompanhar a produção dos hospitais, designadamente:
 - A medição da produção hospitalar;
 - O perfil das admissões.
- Definir e operacionalizar um modelo de avaliação do desempenho hospitalar, mediante a utilização do ajustamento pelo risco, tendo em atenção os resultados destas organizações de saúde no que se refere:
 - À efectividade;
 - À eficiência;
 - A um modelo global/misto.
- Identificar algumas razões justificativas do desempenho dos hospitais, tanto relacionadas com a oferta, especialmente a dimensão da produção e o perfil das admissões, como as respeitantes à procura, tendo especialmente em atenção as características dos doentes, designadamente o tipo e a gravidade da doença;
- Identificar algumas consequências para todos os agentes económicos, dos indicadores de monitorização da produção hospitalar, bem como dos resultados do modelo de avaliação do desempenho hospitalar, tendo em vista a melhoria do financiamento, da regulação, da gestão e do funcionamento dos hospitais.

METODOLOGIA

- População em Estudo e Critérios de Exclusão
- Variáveis em Estudo
- Instrumentos
- Procedimentos para Aplicação do *Software* e Recalibração

4. Metodologia

Na Metodologia são considerados quatro aspectos: a população em estudo e os critérios de exclusão, as variáveis, os instrumentos e os procedimentos utilizados para a identificação da gravidade e para a recalibração dos dados.

Em termos gerais opta-se por utilizar dados de três anos, respeitantes à população internada nos hospitais públicos portugueses (os Açores e Madeira estão excluídos, visto que não utilizam as mesmas fontes de informação utilizadas neste estudo), aos quais foram ainda introduzidos alguns critérios de exclusão.

As variáveis utilizadas respeitam a duas áreas, a produção e o desempenho e a cinco dimensões, medição da produção, perfil das admissões, efectividade dos cuidados prestados, eficiência dos cuidados prestados e desempenho global dos hospitais, pertencendo as duas primeiras à área produção e as três restantes à área desempenho.

As variáveis são definidas conceptual e operacionalmente, sendo genericamente as seguintes:

- Produção
 - Medição da Produção
 - Índice de Casemix (Índice de Casemix dos DRGs, Índice de Casemix do Disease Staging e Índice de Casemix Conjunto)
 - Perfil das Admissões
 - Concentração/diversificação total da população (raridade)
 - Concentração/diversificação da produção a 80%
 - Complexidade dos Casos
 - Gravidade dos Casos
 - Complexidade e Gravidade dos Casos
- Desempenho
 - Efectividade
 - Relação entre a taxa de mortalidade observada e a taxa de mortalidade esperada
 - Eficiência
 - Relação entre a demora média observada e a demora média esperada
 - Desempenho global
 - Média não ponderada da efectividade e da eficiência

Os instrumentos utilizados neste estudo respeitam principalmente a dois sistemas de classificação de doentes – os Diagnosis Related Groups (DRGs) e o Disease Staging – os quais foram anteriormente genericamente caracterizados.

Os procedimentos utilizados no estudo respeitam essencialmente ao seguinte:

- Uniformização das versões dos DRGs, visto que no período em análise existiram duas versões dos DRGs;
- Aplicação do “software” do Disease Staging, para identificar a doença principal, os estadios e os sub-estadios da doença principal e das comorbilidades e as previsões para a mortalidade e para a duração de internamento;
- Recalibração da mortalidade esperada;
- Recalibração da duração de internamento esperada.

O tratamento estatístico foi feito com recurso ao SPSS (versão 11.5), os gráficos e o apuramento de quadros foram executados no Excel (2003) e a edição de texto no Word (2003).

4.1. População em Estudo e Critérios de Exclusão

A população em estudo corresponde aos doentes internados em 88 hospitais públicos do continente português nos anos de 1999, 2000 e 2001, aos quais correspondem os respectivos episódios de internamento.

Neste estudo, os termos “doentes”, “admissões”, “casos” serão utilizados como sinónimos de “episódios de internamento”.

Convém ainda fazer um esclarecimento adicional sobre a semântica da doença. Assim, utiliza-se DRG quando se está a falar dos produtos identificados pelos Diagnosis Related Groups, Doença Principal quando se está a falar dos produtos identificados pelo Disease Staging e de Diagnósticos Principais, quando se está a falar da CID-9-MC.

A utilização dos três anos para se avaliar a actividade e o desempenho dos hospitais segue as indicações de lezzoni e outros (1997c) e O`Muirhearthaigh, Murphy e Moore (2002), entre outros, nas quais se defende que a utilização de três anos em detrimento de um ano constitui uma metodologia mais sólida e segura, visto que desta forma se atenuam eventuais valores extremos.

A fonte de dados corresponde aos resumos informatizados de alta dos hospitais, disponibilizada pela Direcção Geral da Saúde (DGS).

Para os anos em estudo a população alvo é a seguinte:

- 1999 – 935424 episódios de internamento
- 2000 – 941228 episódios de internamento
- 2001 – 943208 episódios de internamento
- Total – 2819860 episódios de internamento

Para se realizar o estudo foram considerados alguns critérios de exclusão, designadamente os “Recém-nascidos e lactentes com afecções do período perinatal” (Neonatologia) e os episódios de internamento cuja saída do hospital não correspondeu a “alta vivo” ou a “óbito” (Episódio Completo), sendo assim retirados os casos correspondentes a transferências para outros hospitais, alta contra parecer médico e seguimento em serviço domiciliário.

Por razões operacionais e que serão explicadas posteriormente existem alguns episódios de internamento que foram incorrectamente classificados pelo “software” do Disease Staging, pelo que foram igualmente retirados da análise.

No Quadro V são apresentadas as populações inicial e final, bem como os episódios eliminados pelos motivos referidos anteriormente.

Quadro V
População Inicial e População Final
Motivos para a Eliminação de Episódios

Anos	População Inicial	Neonatologia	Episódio Completo	“Software”	População Final
1999	935424	102524 (10.96%)	57510 (7.00%)	11909 (1.43%)	775390 (82.89%)
2000	941228	105489 (11.21%)	78528 (9.69%)	25330 (3.03%)	757211 (80.45%)
2001	943208	98234 (10.41%)	62386 (7.49%)	11767 (1.39%)	782588 (82.97%)
Total	2819860	306247 (10.86%)	198424 (8.05%)	49006 (1.95%)	2315189 (82.10%)

A não consideração dos “Recém-nascidos e lactentes com afecções do período perinatal” originou a eliminação de 306427 episódios de internamento, correspondendo a cerca de 10.9% do total dos episódios. Deve evidenciar-se que no ano de 2000 se observou o número mais elevado de eliminações.

A não inclusão de doentes cuja alta não correspondeu a um episódio com tratamento completo originou a eliminação de 198424 doentes (8.05%), mais uma vez com maior expressão em 2000 (9.69%).

Finalmente os erros na programação do Disease Staging (ver sub-capítulo Procedimentos) originaram a eliminação de 49006 episódios de internamento (1.95% do total de episódios), com valores mais elevados também em 2000 (3.03%).

Assim, a população em estudo corresponde a 2315189 episódios de internamento, os quais representam 82.1% da população inicial. O ano de 2000 é o que apresenta menos episódios, o que decorre essencialmente do facto de apresentar sempre eliminações superiores aos dos restantes anos em estudo.

4.2. Variáveis em Estudo

Conforme foi referido na definição dos objectivos do estudo, este centra-se na identificação e operacionalização de indicadores caracterizadores de duas grandes áreas: a Produção e o Desempenho dos hospitais.

Para a Produção dos hospitais são consideradas duas dimensões:

- A Medição da Produção (avaliada pelo Índice de Casemix);
- O Perfil das Admissões (avaliado através da Concentração da Produção, da Complexidade, da Gravidade, e de um indicador combinado da Complexidade e da Gravidade).

No que se refere à avaliação do Desempenho dos hospitais são, consideradas três dimensões:

- A Efectividade dos cuidados prestados;
- A Eficiência dos cuidados prestados;
- O Desempenho Global dos hospitais (Efectividade e Eficiência).

4.2.1. Definições Conceptuais

Produção e Medição da Produção:

“Casemix” – Variedade das situações clínicas dos doentes tratados por cada hospital, organização de saúde ou prestador (Lichtig, 1986).

Índice de Casemix – Valor que expressa a diversidade dos casos tratados em cada hospital (Lichtig, 1986).

Perfil de Admissões:

Concentração da Produção em cada hospital – Valor que expressa a proporção de produtos diferentes tratados em cada hospital (Hornbrook, 1982). É diferente do conceito de concentração da produção entre hospitais, visto que neste indicador se pretende medir a concentração de determinados produtos num número pequeno de hospitais (Schumaker et al, 1979).

Complexidade dos casos – Medida que expressa a quantidade de recursos necessária para tratar determinado caso (Luke, 1979). É diferente do conceito intensidade, visto que nesta perspectiva se pretende medir os recursos necessários para tratar determinado caso por dia de internamento (Luke, 1979). Os casos mais complexos podem estar concentrados em alguns hospitais, atendendo essencialmente às exigências tecnológicas e de recursos humanos que estão associadas a estas situações (Hornbrook, 1982).

Gravidade ou Severidade – Probabilidade de morte ou de falência de um órgão (Thomas, Ashcraft e Zimmermam, 1986; Costa, 1991; lezzoni, 1997b).

Desempenho:

Efectividade – A capacidade de uma intervenção, tratamento ou medicamento melhorar a saúde de uma pessoa ou de uma população, ou ainda, os resultados ou consequências de determinado procedimento ou tecnologia médica quando aplicados na prática (McGuire, Henderson e Mooney, 1988; Pereira, 1993).

Eficiência Técnica – Medida que exprime a relação entre os recursos utilizados e os resultados obtidos. A actividade é eficiente quando se maximizam os resultados para um dado nível de recursos ou se minimizam os recursos para se obter um determinado resultado, como por exemplo, número de consultas por médico ou demora média hospitalar (Pereira, 1993).

4.2.2. Definições Operacionais

Índice de Casemix

O Índice de Casemix é uma medida escalar que sintetiza a diversidade dos produtos hospitalares. Atentas as características, adiante descritas, dos dois sistemas de classificação de doentes (Diagnosis Related Groups e Disease Staging) utilizados neste estudo pode referir-se que os DRGs disponibilizam informações sobre a complexidade dos casos, enquanto que o Disease Staging produz informação sobre a gravidade dos casos. Assim, a utilização conjunta destes dois sistemas de classificação fornece informação sobre a complexidade e a gravidade dos casos tratados.

Neste sentido, são criados três Índices de Casemix:

Índice de Complexidade por hospital

$$\sum_i^n DSdos_i^n * PR_i^n / \sum DSdos$$

Índice de Gravidade por hospital

$\sum_i^n DSdos_i^n * IGD_i^n / \sum DSdos$, sendo o Índice de Gravidade por Doente (IGD) = Gravidade do Doente / Gravidade Média da População

Índice de Casemix Composto por hospital

(Índice de Complexidade + Índice de Gravidade) / 2

Em que:

DSdos – Doentes Saídos;

IGD – Índice de Gravidade do Doente ou Escala de Mortalidade;

IGD_i^n – corresponde ao índice de gravidade i... n. A gravidade é igual à taxa de mortalidade esperada, após recalibração aos dados portugueses;

PR – Peso Relativo de cada DRG. Expressa a complexidade de cada DRG. Neste estudo foram utilizados os Pesos Relativos constantes da Portaria nº189/2001 de 9 de Março.

A este propósito é conveniente fazer alguns comentários suplementares.

Em Portugal, para os DRGs o Índice de Casemix é calculado em função da duração de internamento dos doentes em cada DRG, dando-se ponderações diferentes para doentes com duração normal, curta duração, longa duração e duração prolongada.

Para tal foi criada a noção de doente equivalente, de acordo com a seguinte metodologia para a ponderação de cada doente (Bentes et al, 1996):

- “Doentes Normais” = 1;
- “Doentes de Curta Duração” = Duração de Internamento / Demora Média do DRG;
- “Doentes de Longa Duração” = $1 + 0.6 * (Duração de Internamento - (Limiar Superior + 1) / Demora Média do DRG)$;
- “Doentes com Duração Prolongada” = $1 + 0.6 * (Limiar Máximo - (Limiar Superior + 1) / Demora Média do DRG)$.

Para qualquer dos grupos de doentes, a consideração de valores normais, de curta duração, bem como os Limiares Superiores e Máximos estão definidos nas Tabelas de Preços dos DRGs, publicadas em 9 Portarias (No Anexo II pode ser encontrada a Portaria nº 189/2001 de 9 de Março, a qual foi utilizada neste estudo).

Assim, o total de doentes equivalentes em cada DRG e hospital é diferente do total de doentes tratados, visto que o primeiro indicador resulta directamente da proporção de casos em cada uma das categorias de doentes anteriormente referidas.

Contudo, esta metodologia não parece ser a mais correcta, visto que existe uma dupla penalização dos hospitais mais eficientes, em primeiro lugar, porque os doentes de estadia mais curta, desde que estejam bem classificados e para um nível de qualidade aceitável, são menos valorizados; em segundo lugar, porque um hospital ao tratar uma proporção menor de doentes com estadias superiores ao limiar superior vê igualmente reduzido o número de doentes com coeficiente superior a 1.

Neste sentido, para se minimizar o efeito da neutralidade económica sempre presente nos DRGs (Costa, 1994), adota-se a fórmula bruta para se calcular o Índice de Casemix segundo os DRGs.

Para o Disease Staging a gravidade é calculada em função da mortalidade prevista para cada doente, após recalibração aos dados portugueses (ver no capítulo Metodologia, o sub-capítulo Procedimentos).

Assim, foi construída uma escala de mortalidade que pode ser calculada por doença (produto), ou para o total dos doentes.

A principal vantagem da primeira alternativa é a maior especificidade que a medida tem, contudo esta deve ser contrastada com a vantagem que existe de um comparador global, o qual permite uma avaliação mais directa de todos os doentes. É, essencialmente por esta última razão que, neste estudo se vai utilizar a escala de mortalidade calculada através do valor esperado médio, como “proxy” da gravidade de cada doente.

No entanto, para que estes valores esperados de mortalidade possam ser utilizados como “proxy” da gravidade dos doentes e concomitantemente do “casemix” é necessário avaliar o poder discriminativo, por hospital, da mortalidade esperada. Visto que, somente se podem utilizar os valores médios da taxa de mortalidade esperada como representativos da gravidade quando o modelo apresenta um bom poder discriminativo (Ash e Shwartz, 1997). No Quadro VI são apresentadas por hospital os valores da estatística “c”.

Os dados permitem concluir que, para a globalidade dos hospitais o novo valor esperado para a taxa de mortalidade apresenta uma boa discriminação, com um valor máximo da estatística “c” de 0.965 (hospital 20) e um valor mínimo de 0.725 (hospital 18). Se ainda se considerar que em 50% dos hospitais a estatística “c” apresenta valores iguais ou superiores a 0.9 e que, em cerca de

47% dos hospitais esta estatística se encontra entre 0.8 e 0.899, mais facilmente se compreende que existe discriminação na taxa de mortalidade esperada.

Este facto permite que estes valores da taxa de mortalidade esperada possam ser utilizados como um “proxy” da gravidade média dos doentes tratados em cada hospital e consequentemente, numa fase posterior deste estudo, apurar se este tipo de indicador introduz ou não novos elementos na avaliação da produção, da efectividade e da eficiência dos hospitais.

Quadro VI
Discriminação (estatística “c”) da Taxa de Mortalidade Esperada
por Hospital, após Recalibração

Hospital	Est “c”	Hospital	Est “c”	Hospital	Est “c”
1	0.869	31	0.878	60	0.906
2	0.904	32	0.913	61	0.910
3	0.893	33	0.903	62	0.899
4	0.883	34	0.894	63	0.917
5	0.880	35	0.878	64	0.914
6	0.860	36	0.902	65	0.854
7	0.880	37	0.892	66	0.841
8	0.908	38	0.904	67	0.857
9	0.919	39	0.918	68	0.857
10	0.905	40	0.872	69	0.924
11	0.791	41	0.907	70	0.858
12	0.850	42	0.925	71	0.871
13	0.873	43	0.864	72	0.880
14	0.841	44	0.887	73	0.763
15	0.881	45	0.923	74	0.887
16	0.901	46	0.899	75	0.842
17	0.941	47	0.900	76	0.904
18	0.725	48	0.903	77	0.879
19	0.938	49	0.904	78	0.877
20	0.965	50	0.870	79	0.894
21	0.940	51	0.875	80	0.899
22	0.904	52	0.880	81	0.816
23	0.923	53	0.904	82	0.926
24	0.911	54	0.899	83	0.846
25	0.903	55	0.908	84	0.861
26	0.913	56	0.912	85	0.926
27	0.912	57	0.880	86	0.912
28	0.929	58	0.932	87	0.915
29	0.914	59	0.907	88	0.914
30	0.886				

A principal razão para se compararem/usarem os dois indicadores de “casemix” (utilizando os DRGs ou o Disease Staging) deriva do facto de se presumir que os

sistemas de classificação de doentes utilizados estarem a medir e a incorporar dimensões diferentes do processo produtivo.

Enquanto os DRGs valorizam mais a utilização e o consumo de recursos, o Disease Staging dá prioridade às características dos doentes que têm valor prognóstico em relação aos resultados de saúde. Por outro lado, o modelo empírico dos DRGs e, naturalmente os respectivos pesos relativos, baseia-se essencialmente no que foi feito pelos hospitais para tratar cada “caso”, e o Disease Staging, pode exprimir, pelo menos de forma mediata, aquilo que seria necessário fazer para se tratar cada “caso”.

Assim, as questões mais relevantes em análise estão relacionadas com o facto de se saber se as duas dimensões facultadas por ambos os sistemas de classificação de doentes, níveis de utilização de recursos no caso dos DRGs e “carga da doença” com o Disease Staging, são ou não mutuamente exclusivas e se disponibilizam ou não informações diferentes sobre o “casemix” dos hospitais.

Concentração da Produção em cada hospital

As principais razões para se estudar a concentração da produção são de dupla natureza: em primeiro lugar, avaliar se existe um comportamento distinto por parte dos hospitais, e na eventualidade de tal ocorrer, se estes comportamentos estão relacionados com algumas características dos hospitais; em segundo lugar, porque existe alguma evidência na literatura económica que a concentração/especialização da produção pode ter alguma relação directa com o desempenho dos hospitais, tanto em termos de eficiência, como da efectividade dos cuidados prestados (Horn e Schumaker, 1979; Watts e Klastorin, 1980; Goldfarb, Hornbrook e Higgins, 1983; Hornbrook e Monheit, 1995; McGuire e Williams, 1986; Farley e Hogan, 1990; Shwartz, Ash e Iezzoni, 1997; Gowrisankaran e Town, 2003 e Shen, 2003).

Serão utilizadas as informações disponibilizadas pelos Diagnosis Related Groups (DRGs) e pelo Disease Staging, tendo essencialmente em vista a discussão se ambos os sistemas de classificação estão a produzir os mesmos resultados e, na eventualidade de tal não acontecer, quais as principais razões para esta situação e qual a melhor forma de a contornar.

Para a concentração da produção irão ser analisadas duas perspectivas:

- Número de produtos diferentes responsáveis pelo total da produção;
- Número de produtos responsável por 80% da produção.

Estes indicadores são construídos através da divisão do número de produtos do hospital pelo total de produtos nos hospitais portugueses, para a primeira medida e pelo número de produtos responsáveis por 80% da produção pelo total de produtos tratados em cada hospital, para a última medida. Como se referiu os

produtos são identificados segundo a metodologia dos Diagnosis Related Groups e do Disease Staging.

A fórmula de cálculo deste(s) indicador(es), expressa em percentagem, permite concluir que quanto menor for seu o valor maior é a concentração da produção em cada hospital, ou ainda o seu grau de especialização na prestação de cuidados de saúde ou vice-versa.

Complexidade dos Casos

Para se avaliar a complexidade dos casos por hospital serão utilizados três indicadores alternativos:

- A identificação de Níveis de Complexidade
- O apuramento da Complexidade Global
- O Índice de Casemix dos DRGs por hospital (já descrito)

A metodologia para apuramento do Índice de Casemix foi descrita anteriormente, pelo que se passa a descrever a forma de cálculo dos outros dois indicadores: (1) Identificação de Níveis de Complexidade e (2) Apuramento da Complexidade Global dos hospitais.

Para se operacionalizar o primeiro aspecto são utilizados os pesos relativos por DRG em vigor em 2001 (Portaria nº 189/2001 de 9 de Março) para se definirem Níveis de Complexidade dos produtos dos hospitais.

Para tal foram definidos três níveis de complexidade, de acordo com a seguinte metodologia:

- Identificação dos níveis de complexidade através dos percentis 33.3 e 66.7, para os pesos relativos;
- Atribuição do nível de complexidade 1 a todos os DRGs cujos pesos relativos estejam incluídos até ao percentil 33.3. Com o nível 2 de complexidade são classificados os DRGs com pesos relativos entre os percentis 33.3 e 66.7. Finalmente para o nível 3 são considerados os DRGs cujos pesos relativos sejam superiores ao percentil 66.7.

Assim, os Níveis de Complexidade são os seguintes:

- Nível 1 – DRGs com pesos relativos até 0.82 (162 DRGs);
- Nível 2 – DRGs com pesos relativos entre 0.84 e 1.53 (164 DRGs);
- Nível 3 – DRGs com pesos relativos superiores a 1.54 (163 DRGs).

Para avaliar a Complexidade Global (medida não ponderada) vai ser utilizada uma medida normalizada (“z score”) de acordo com a seguinte equação:

“Z score Admissão” = (Número de Doentes tratados no Nível 1 de Complexidade – Número de Doentes Tratados no Nível 2 de Complexidade – Número de Doentes Tratados no Nível 3 de Complexidade) / Desvio Padrão dos Doentes Tratados no Nível 1 de Complexidade.

Gravidade dos Casos

Para se avaliar a gravidade dos casos por hospital serão utilizadas duas medidas alternativas:

- A identificação de Estádios de Gravidade
- O apuramento da Gravidade Global
- O Índice de Casemix do Disease Staging por hospital (já descrito)

A metodologia para apuramento do Índice de Casemix foi descrita anteriormente, pelo que se passa a descrever a forma de cálculo dos outros dois indicadores: (1) Identificação de Níveis de Gravidade e (2) Apuramento da Gravidade Global.

Para se identificarem os Estádios de Gravidade são utilizados os estádios de severidade do Disease Staging, correspondendo genericamente:

- Estadio 1 – Doentes menos graves;
- Estadio 2 – Doentes com gravidade intermédia;
- Estadio 3 – Doentes mais graves.

Para avaliar a Gravidade Global vai ser utilizada uma medida normalizada (“z score”) de acordo com a seguinte equação:

“Z score Admissão” = (Número de Doentes tratados no Estadio 1 – Número de Doentes Tratados no Estadio 2 – Número de Doentes Tratados no Estadio 3) / Desvio Padrão dos Doentes Tratados no Estadio 1.

Efectividade

A efectividade por hospital vai ser medida pela relação entre a Taxa de Mortalidade Observada e a Taxa de Mortalidade Esperada.

Os procedimentos para se apurar a Taxa de Mortalidade Esperada são descritos no sub-capítulo “Procedimentos para aplicação do “software” e de recalibração”.

Para se apurar aquela relação existem duas alternativas: (1) Padronização Indirecta e Padronização Directa (Daley, 1997; Iezzoni, 1997c e Schwartz, Ash e Iezzoni, 1997).

Padronização Indirecta (Shwartz, Ash e Iezzoni, 1997 e O'Muircheartaigh, Murphy e Moore, 2002).

Efectividade = Taxa de Mortalidade Observada / Taxa de Mortalidade Esperada.

Padronização Directa (Daley, 1997 e Iezzoni, 1997c).

Efectividade ("z score") = (Número de Óbitos Observados – Número de Óbitos Esperados) / Desvio Padrão dos Óbitos Observados.

Eficiência

A eficiência por hospital vai ser medida pela relação entre a Demora Média Observada e a Demora Média Esperada. Os procedimentos para se apurar a Demora Média Esperada são descritos no sub-capítulo Procedimentos.

À semelhança do referido para a efectividade, para se apurar esta relação existem duas alternativas: (1) Padronização Indirecta e Padronização Directa.

Padronização Indirecta (HCIA, 1999).

Eficiência = Demora Média Observada / Demora Média Esperada.

Padronização Directa (Daley, 1997 e Iezzoni, 1997c).

Eficiência ("z score") = (Duração de Internamento Observada – Duração de Internamento Esperada) / Desvio Padrão da Duração de Internamento Observada.

Desempenho Global

Padronização Indirecta

- Cenário A – valorização idêntica da efectividade e da eficiência
Indicador conjunto = (Ordenação da efectividade + Ordenação da eficiência) / 2;
- Cenário B – maior valorização da efectividade
Indicador conjunto = (2 * Ordenação da efectividade + Ordenação da eficiência) / 3;
- Cenário C – maior valorização da eficiência
Indicador conjunto = (Ordenação da efectividade + 2 * Ordenação da eficiência) / 3.

Padronização Directa

- Cenário I – valorização idêntica da efectividade e da eficiência
 $\text{Indicador conjunto} = (\text{"z score" da efectividade} + \text{"z score" da eficiência}) / 2;$
- Cenário II – maior valorização da efectividade
 $\text{Indicador conjunto} = (2 * \text{"z score" da efectividade} + \text{"z score" da eficiência}) / 3;$
- Cenário III – maior valorização da eficiência
 $\text{Indicador conjunto} = (\text{"z score" da efectividade} + 2 * \text{"z score" da eficiência}) / 3.$

4. 3. Instrumentos

Conforme foi referido, o estudo centra-se somente no Internamento e envolve duas áreas – Produção e Desempenho.

Independentemente da perspectiva de avaliação, o ajustamento pelo risco é necessário para se avaliar a actividade das organizações de saúde. Isto porque é necessário medir e controlar as características dos doentes que podem influenciar os resultados de saúde.

Em termos genéricos a principal finalidade do ajustamento pelo risco é a de controlar os factores que os doentes apresentam ao contactar uma determinada organização de saúde que podem afectar a sua probabilidade de obterem um bom ou um mau resultado. Assim, o “rationale” para o ajustamento pelo risco consiste na eliminação dos factores que podem conduzir a diferentes resultados de saúde, essencialmente aqueles que estão associados aos doentes.

Em Portugal, para os hospitais e para a finalidade que se pretende, a informação existente baseia-se nos resumos de alta e consequentemente pode ser qualificada como do tipo administrativo, permitindo igualmente o desenvolvimento de sistemas retrospectivos de ajustamento pelo risco.

As principais excepções são encontradas em algumas Unidades de Cuidados Intensivos (UCIs), nas quais são utilizados sistemas prospectivos de medição da severidade do estado do doente, como o APACHE, o SAPS e o MPM (Carneiro, 1994).

Contudo, atendendo a que estas últimas modalidades não estão generalizadas a todos os serviços do hospital e a que ainda não estão totalmente validadas para outro tipo de doentes para além daqueles que são tratados nas UCIs, interessa agora caracterizar o sistema de informação existente para os doentes internados.

A folha de “admissão e de alta” permite identificar os seguintes elementos (IGIF):

- Identificação do doente
 - Número de Processo Clínico
 - Sexo
 - Data de Nascimento
 - Entidade Responsável pelo Financiamento
 - Número de Beneficiário
 - Residência (Distrito, Concelho e Freguesia)
- Natureza da Admissão
 - Admissão Programada
 - Admissão Não Programada

- Transferências entre Serviços (20 campos, para 1ª, 2ª, 3ª...19ª e último serviço)
 - Código do Serviço
 - Designação do Serviço
 - Data de Admissão
 - Data de Alta
- Destino Após a Alta
 - Para o Domicílio
 - Hospital de Dia
 - Consulta Externa do Hospital
 - Consulta Externa de Outro Hospital
 - Centro de Saúde/Médico de Família
 - Sem outra Especificação
 - Para outra Instituição com Internamento
 - Hospital do SNS
 - Centro de Saúde
 - Outro Hospital
 - Serviço Domiciliário
 - Saído Contra Parecer Médico
 - Falecido
- Transferência entre Hospitais
 - Transferido Para:
 - Realização de Exames
 - Seguimento
 - Por Falta de Recursos
 - Tratamento de Condição Associada
 - Transferido De
- Diagnóstico de Admissão
- Morfologia Tumoral
- Peso à Nascimento
- Diagnósticos – até 20 e com indicação do Diagnóstico Principal
- Causa Externa de Lesão/Efeito Adverso – até 20
- Procedimentos – até 20
- Data da primeira intervenção cirúrgica
- Número de Dias em Unidades de Cuidados Intensivos
- Identificação do Médico Responsável pelo Tratamento
- Identificação do Médico Codificador

Para a codificação dos dados é utilizada a CID-9-MC, sendo ainda de referir que a possibilidade de se utilizarem até 20 campos para os diagnósticos, para as causas externas de lesão/efeitos adversos e para os procedimentos, minimiza alguns dos problemas referidos por Iezzoni e outros (1992c), no que respeita à utilização de dados administrativos para o desenvolvimento de sistemas de ajustamento pelo risco, embora deva somente ser considerada uma condição

necessária, mas não suficiente para melhorar as características e propriedades dos referidos sistemas.

Neste sentido, atentos os aspectos referidos no Enquadramento Teórico e Questões Práticas, embora existisse a possibilidade de recorrer a versões dos DRGs que incorporem a gravidade, designadamente os All Patient Refined DRGs ou os International Refined DRGs, neste estudo serão somente utilizadas a versão dos DRGs em vigor em Portugal e o Disease Staging.

Recorde-se ainda que enquanto os DRGS são um sistema de classificação de doentes que identifica produtos cujos diagnósticos estão relacionados e com consumo homogéneo de recursos, constituindo uma metodologia que identifica a complexidade dos produtos (Fetter et al, 1980 e McGuire, 1991) o Disease Staging identifica produtos em função de um modelo biológico de progressão da doença em função do respectivo prognóstico, pelo que disponibiliza informação sobre a gravidade do doente (Thomas, Ashcraft e Zimmerman, 1986 e Gonnella, Louis e Gozum, 1999).

4. 4. Procedimentos para Aplicação do “Software” e Recalibração

Tendo em vista o apuramento das variáveis referidas e recorrendo aos dois sistemas de classificação de doentes descritos foram realizados os seguintes procedimentos para aplicação do “software” e recalibração:

- Uniformização das versões dos DRGs;
- Aplicação do “software” do Disease Staging;
- Recalibração da mortalidade esperada;
- Recalibração da duração de internamento esperada.

Uniformização das versões dos DRGs

Em Portugal até ao presente momento, foram utilizadas três versões do sistema de DRGs desenvolvido pela HCFA, as Versões 6, 10 e 16, as quais entraram em vigor, respectivamente em 1990, 1996 e 2001, embora as mesmas tivessem sido introduzidas nos EUA em 1988, 1990 e 1996 (Santana, 2003).

Contudo, o projecto de introdução dos DRGs teve início em 1984, tendo em vista estudar a viabilidade da sua implementação em Portugal (Urbano e Bentes, 1990). Atendendo aos resultados obtidos, iniciou-se em 1987 o processo conducente à concepção e implementação de um sistema prospectivo de financiamento hospitalar (Rodrigues, 2002).

Segundo os seus responsáveis a estratégia adoptada para o financiamento hospitalar foi faseada, começando por ser aplicada a terceiros pagadores e posteriormente aplicada ao Serviço Nacional de Saúde, através da atribuição de uma percentagem progressivamente crescente do valor correspondente aos DRGs em relação aos orçamentos dos hospitais (Bentes et al, 1996; Dismuke, 1996 e Rodrigues, 2002).

Atendendo a que neste estudo será utilizado o sistema desenvolvido pela HCFA e que no período em análise (1999 a 2001) existiram duas versões, a versão 10 até ao ano 2000 e a versão 16 a partir de 2001 é ainda conveniente fazer algumas observações suplementares.

As principais diferenças entre as Versões 10 e 16 dos DRGs respeitam ao número de categorias de DRG, bem como no número de DRGs, à existência de uma Pré-Grande Categoria Diagnóstica para Transplantes na Versão 16 e à introdução de uma Grande Categoria Diagnóstica (GCD) para o VIH (Santana, 2003).

Deve ainda referir-se que, apesar de terem existido estas duas versões de DRGs, neste estudo são somente utilizados dados referentes à versão 16. Para que tal fosse concretizado adoptou-se o seguinte procedimento:

- Para 2001 foram utilizados os dados produzidos pelos DRGs;
- Para 2000, embora a versão em vigor fosse a 10, o facto dos dados produzidos pelos DRGs incluírem igualmente a versão 16 permitiu que fosse utilizada esta informação;
- Para 1999 os dados produzidos pelos DRGs incluíam somente a versão 10, pelo que foram utilizados os dados derivados do “software” do Disease Staging para se recodificarem os doentes na Versão 16.

A este propósito convém referir que pese embora o facto de o “software” Disease Staging disponibilizar informação sobre o DRG atribuído a cada doente, os valores não coincidem perfeitamente com os facultados pelo “software” dos DRGs (em função dos dados disponibilizados pela Direcção Geral de Saúde). Este facto foi confirmado em 2001, onde se observou que em cerca de 2% dos doentes foram atribuídos DRGs diferentes (22482 doentes).

Atendendo à dimensão da população em estudo (2891260 doentes) pode considerar-se que a eventual atribuição de um DRG diferente aos doentes não originou desvios e/ou enviesamentos no estudo. De facto, recorde-se que tal somente ocorreu em 1999, pelo que somente poderá originar um desvio potencial em 0.8% do total de episódios.

Aplicação do “software” do Disease Staging

O “software” do Disease Staging é propriedade da MEDSTAT, uma empresa do grupo THOMSON, tendo neste estudo sido utilizada a versão 4.12 (MEDSTAT, 2001).

Este “software” está programado para utilizar, tanto a Classificação Internacional das Doenças, 9ª Revisão; Modificações Clínicas (CID-9-MC), como a 10ª Revisão desta mesma classificação de doenças, pelo que não existiram quaisquer problemas para a sua utilização nos dados portugueses, visto que se encontra em vigor a primeira versão referida.

Atendendo a que o “software” trabalha em aplicações do tipo “txt”, foram feitos os seguintes procedimentos de programação:

1. Conversão da base de dados dos resumos de alta dos hospitais de ficheiros do tipo Acess ou Dbase para ficheiros do SPSS;
2. Conversão dos ficheiros do SPSS para ficheiros “txt”;
3. Programação do “software” do Disease Staging:

- Identificação do ficheiro a programar, do ficheiro de “output” e do ficheiro de erros;
 - Identificação, por episódio, dos campos dos “inputs” (sexo, idade, tipo de admissão, destino após a alta, diagnósticos e procedimentos), início do campo, dimensão do campo e códigos dos campos (quando necessário). Por exemplo, para os dados portugueses somente foi necessário identificar os códigos para o sexo, para o tipo de admissão e para o destino após a alta;
 - Identificação, por episódio, dos campos de “outputs”, igualmente com referência ao início do campo e da respectiva duração:
 - DRG
 - Doença Principal
 - Comorbilidades
 - Estadio e Sub-Estadios do Diagnóstico Principal
 - Estadio e Sub-Estadios das Comorbilidades
 - Custo Previsto (a)
 - Duração de Internamento Prevista
 - Mortalidade Prevista

(a) atendendo a que nos hospitais portugueses não existem custos por doente, esta variável não foi utilizada no estudo.
4. Conversão do ficheiro de “output” do “software” do Disease Staging para um ficheiro SPSS.

Conforme foi referido a aplicação deste “software” originou a perda de 49006 doentes, cerca de 2% do total da população em estudo, com maior destaque no ano de 2000 (25330 episódios) e menor expressão em 2001 (11776 episódios).

Recalibração da mortalidade esperada

Nesta fase do estudo irão somente ser descritos os procedimentos relativos à definição de estimativas para os valores esperados de mortalidade. Tendo em atenção que os procedimentos para a recalibração devem aproximar-se do modelo original, em seguida serão descritos os passos adoptados pela MEDSTAT. Assim, na generalidade, a metodologia utilizada pelo Disease Staging foi a seguinte (MEDSTAT, 2001):

- Separação dos casos cirúrgicos e médicos para a definição de valores esperados de mortalidade. Para este efeito embora a versão em vigor em Portugal dos Diagnosis Related Groups (DRGs) identifique casos cirúrgicos, médicos e outros, para efeitos de mortalidade são considerados casos cirúrgicos e não cirúrgicos;
- Para os casos médicos são consideradas duas situações:
 - Doenças com menos de 300 admissões, doenças com taxa de mortalidade ≤ 0.005 e doenças com óbitos ≤ 25 – os valores esperados e observados admitem-se iguais;

- Doenças com 300 ou mais admissões – desenvolveu-se um Peso de Comorbilidade, o qual é uma função da importância relativa de cada comorbilidade em relação à probabilidade de morte de acordo com a seguinte equação: idade igual a zero + idade + idade ao quadrado + raiz quadrada da idade + sexo + sub-estadio. Em seguida calculou-se o Modelo de Mortalidade de acordo com a seguinte equação: sub-estadio + peso de comorbilidade + peso de comorbilidade ao quadrado + raiz quadrada do peso de comorbilidade + idade + idade ao quadrado + raiz quadrada da idade + sexo + admissão electiva + transferência + admissão urgente + número de diagnósticos.
- Para os casos cirúrgicos são igualmente consideradas duas situações:
 - DRGs com menos de 300 casos e DRGs com taxas de mortalidade ≤ 0.01 – o valor esperado é igual ao observado;
 - DRGs com 300 ou mais casos – metodologia semelhante à descrita para os casos médicos, com a particularidade de os cálculos serem efectuados, tanto ao nível do DRG, como do diagnóstico principal.

NOTA: Idade = 0 é uma variável “dummy” para distinguir os recém nascidos dos outros doentes, assumindo o valor 1 quando a idade é igual a zero e o valor 0 quando a idade > 0.

Em relação às comorbilidades foram realizadas 527 regressões logísticas, quase todas com significância estatística, evidenciando ainda que a utilização do Disease Staging apresenta melhores resultados que os disponibilizados pelo Índice de Comorbilidade de Charlson e que o modelo utilizado é melhor que a utilização de índices de comorbilidades essencialmente baseados na identificação de um determinado número de situações de saúde (Houchens, 2002).

Finalmente, os autores referem ainda que a discriminação do modelo, medida pela estatística “c” é muito boa, visto que o valor apurado é de 0.944 (MEDSTAT, 2001).

Atendendo ao referido anteriormente, interessa agora passar à avaliação do desempenho do Disease Staging quando aplicado aos dados portugueses.

No Quadro VII são apresentados os valores globais e por ano.

Conforme se observa o Disease Staging não está bem adaptado à realidade portuguesa, visto que o valor mínimo da razão entre as taxas de mortalidade observadas e esperadas é de 2.56. Para se considerar o modelo bem calibrado esta razão deveria estar muito próximo de 1.

Quadro VII
Taxas de Mortalidade Hospitalar Observada e Esperada
1999, 2000, 2001 e Total do Período

Ano	Taxa de Mortalidade Observada %	Taxa de Mortalidade Esperada %	Razão
1999	4.58	1.57	2.92
2000	4.67	1.70	2.75
2001	4.69	1.83	2.56
Período	4.65	1.70	2.74

Esta situação parece confirmar os resultados dos estudos de Justice, Covinsky e Berlin (1999), Flanders e outros (1999) e do artigo de Iezzoni (1999), no qual a autora baseando-se naqueles estudos refere que sempre que se transportam resultados de medidas estatísticas para outras populações, tanto em termos espaciais, como temporais, os analistas devem observar o princípio “caveat emptor”. Ou seja, muito dificilmente o modelo aplicado a diferentes populações apresentará resultados precisos. A autora refere ainda que estes problemas de calibração podem ser solucionados com recurso a técnicas estatísticas como a regressão logística.

Por outro lado, somente a título exemplificativo, deve referir-se o facto de a taxa de mortalidade esperada quando se utiliza o Disease Staging, ser progressivamente superior em cada ano do período em estudo. Este facto, pode ocorrer por acaso, mas quando se tem presente o defendido por Iezzoni e outros (1992c) em relação à importância da codificação e se sabe igualmente que 2001 é o único ano em que as bases de dados utilizadas disponibilizam informação até 19 diagnósticos secundários, devem ser desenvolvidos outros estudos para se avaliar a importância deste factor e deve ainda ser enfatizada a importância de se realizarem estudos de fiabilidade sobre a política de codificação de cada hospital.

Estas discrepâncias encontradas para os valores globais entre taxas de mortalidade observadas e esperadas são igualmente encontrados quando se desagrega a análise, por região de saúde e por tipo de hospital segundo a classificação do IGIF (ver Quadro VIII).

Por tipo de hospital observam-se igualmente grandes diferenças entre taxas de mortalidade observadas e esperadas, com maior expressão nos hospitais do tipo IV e III, embora nos de tipo II a razão entre estes indicadores esteja muito próxima de 1.

Quadro VIII
Taxas de Mortalidade, Observada e Esperada
Por Tipo de Hospital e Por Região

Tipo Hospital	Taxa Mortalidade Observada	Taxa Mortalidade Esperada	Razão	Região	Taxa Mortalidade Observada	Taxa Mortalidade Esperada	Razão
I	4.96	2.10	2.36	I	6.02	1.69	3.56
II	0.18	0.17	1.06	II	5.51	1.41	3.91
III	4.53	1.50	3.02	III	4.07	1.64	2.48
IV	5.64	1.68	3.36	IV	5.51	1.91	2.88
V	6.52	2.55	2.56	V	3.95	1.54	2.56

No que se refere às Regiões de Saúde os valores observados para as taxas de mortalidade são sempre superiores aos dos valores esperados, embora com maior expressão nas Regiões II e I.

Embora com intensidades diferentes, estas discrepâncias entre os valores médios das taxas de mortalidade observadas e esperadas, não permite a utilização directa dos resultados do “software” do Disease Staging à realidade portuguesa.

Neste sentido, torna-se imprescindível proceder à recalibração do Disease Staging aos dados portugueses seguindo a metodologia e as orientações definidas por Justice, Covinsky e Berlin (1999), Flanders e outros (1999) e Iezzoni (1999), essencialmente utilizando técnicas de regressão logística.

Para se recalibrarem os dados da mortalidade esperada, decorrente da aplicação do “software” do Disease Staging à realidade portuguesa será utilizada a regressão logística com a seguinte equação (MEDSTAT, 2002):

$$Y = a + b * \text{logit}(p)$$

Em que

Y=0 quando alta vivo e Y=1 quando ocorreu um óbito

e

$$\text{logit}(p) = \ln(p / (1-p))$$

Ln é o logaritmo natural

e

p é igual à mortalidade prevista pelo Disease Staging

Conforme foi referido anteriormente para se utilizarem os valores previstos do modelo é necessário avaliar o seu ajustamento. Para tal existem duas técnicas,

a calibração, medida pelo teste de Hosmer-Lemeshow e a discriminação, avaliada pela estatística “c”.

Foi igualmente discutida a importância de cada uma destas técnicas, sendo contudo de evidenciar que para a situação em análise a calibração é o aspecto determinante, essencialmente porque se pretendem comparar valores observados e esperados (Lemeshow e Hosmer, 1992 e Ash e Schwartz, 1997). Apesar destes argumentos, para cada uma das análises de regressão logística efectuada serão apresentados os valores da calibração e da discriminação.

Quando as taxas de mortalidade observadas são muito baixas, serão utilizados os novos valores previstos, independentemente do ajustamento do modelo.

Após a realização das regressões logísticas o novo valor previsto para a mortalidade é dado pela equação seguinte (MEDSTAT, 2002):

$$\text{Nemt} = 1 / (1 + \exp(-a - b * \text{logit}(p)))$$

Neste sentido, à semelhança da versão original do “software” do Disease Staging utilizam-se procedimentos distintos para casos cirúrgicos e para casos médicos.

Para se proceder à recalibração adoptou-se uma metodologia iterativa:

- 1º Passo – Regressão Logística entre a mortalidade observada e os valores previstos pelo Disease Staging;
- 2º Passo – Regressão Logística entre a mortalidade observada e o novo valor previsto para a mortalidade (decorrente do 1º Passo da regressão logística).

1º Passo

Para os casos cirúrgicos seguiu-se a seguinte metodologia:

1. Regressão por Grandes Categorias de Diagnóstico (GCDs) (ver Anexo II) – C1
2. Se não existir calibração, introduzir variável idade – C2
3. Se não existir calibração, fazer regressão por DRG – C3
4. Se não existir calibração, introduzir variável idade – C4
5. Se não existir calibração, fazer regressão por Categoria Principal de Doença (PDXCAT) – C5
6. Se não existir calibração, introduzir variável idade – C6
7. Casos com baixas taxas de mortalidade (se as taxas de mortalidade observadas forem muito baixas, serão utilizados os novos valores previstos, independentemente do ajustamento do modelo) – C7

8. Casos com frequência baixa (menos de 10 doentes tratados) ou sem óbitos – igual ao valor previsto do Modelo Disease Staging – C8

Para os casos médicos seguiu-se a seguinte metodologia:

1. Regressão pelos Grandes Agrupamentos de Doença (GADs) (ver Anexo I) – M1
2. Se não existir calibração, introduzir variável Categoria Principal de Doença (PDXCAT) – M2
3. Se não existir calibração, introduzir variável idade – M3
4. Se não existir calibração, introduzir variáveis Categoria Principal de Doença e idade – M4
5. Se não existir calibração, utilizar Categoria Principal de Doença – M5
6. Se não existir calibração, introduzir variável idade – M6
7. Se não existir calibração, introduzir variáveis Categoria Principal de Doença e idade – M7
8. Se não existir calibração, utilizar diagnóstico principal (3 dígitos da Classificação Internacional das Doenças, 9ª Revisão) – M8
9. Se não existir calibração, introduzir variável idade – M9
10. Se não existir calibração, introduzir variável diagnóstico principal desagregado (a 4 ou 5 dígitos, da mesma classificação de doenças) – M10
11. Se não existir calibração, introduzir variáveis diagnóstico principal e idade – M11
12. Se não existir calibração, desagregar diagnóstico principal (a 4 ou a 5 dígitos) – M12
13. Se não existir calibração, introduzir variável idade – M13
14. Casos com baixas taxas de mortalidade (se as taxas de mortalidade observadas forem muito baixas, serão utilizados os novos valores previstos, independentemente do ajustamento do modelo) – M14
15. Casos sem calibração (ver descrição na página 36) – M15

Os critérios e resultados da aplicação da técnica de regressão logística para os casos cirúrgicos são apresentados no Quadro IX.

A avaliação dos dados do Quadro IX permite concluir que na grande maioria das situações o modelo de regressão logística mais frequente respeita à consideração de dados por doença [cerca de 81% (C1+C3+C5)]. Destas a mais representativa foi a que utiliza como critério de escolha o DRG (cerca de 66%).

Por outro lado, em três situações utilizou-se o valor previsto do Disease Staging, mais precisamente nos DRGs 6 e 163, onde não se observou nenhum óbito e no DRG 495, onde somente foi tratado um doente. A estas três situações corresponderam 21690 episódios de internamento (2.42% do total dos casos cirúrgicos e 0.94 to total de episódios).

Quadro IX
Critérios para a Recalibração do Disease Staging
Casos Cirúrgicos

Critério de Regressão	Número de Ocorrências	% de Ocorrências
C1	8	9.09
C2	2	2.27
C3	58	65.91
C4	6	6.82
C5	5	5.68
C6	0	0.0
C7	6	6.82
C8	3	3.41
Total	88	100.00

Em 6 situações (cerca de 7%) foi utilizado o novo valor previsto da regressão logística, embora sem calibração, visto estar-se na presença de doenças ou grupos de doenças com baixa mortalidade (ver Quadro X).

Quadro X
Situações sem Calibração
Casos Cirúrgicos

Situações	Doentes Tratados	Taxa de Mortalidade
GCD 3	52585	0.15
GCD 8	142804	0.90
GCD 9 (excepto DRG 263)	66539	0.16
GCD 10 (excepto DRG 285)	10425	0.48
GCD 12	34034	0.00
GCD 13 (excepto DRGs 353, 357 e 365)	81111	0.08
Total	387498	0.43

Os casos referidos no Quadro X representam cerca de 17% do total de episódios de internamento e de 43% dos casos cirúrgicos, mas atendendo às baixas taxas de mortalidade observadas, a inexistência de um bom valor para o teste de Hosmer-Lemeshow não põe em causa a adaptação ou a definição de novos valores previstos da utilização do Disease Staging.

Os critérios e resultados da aplicação da regressão logística para os casos médicos são apresentados no Quadro XI.

Os modelos de regressão logística mais frequentes respeitam:

- Ao modelo M12, ou seja a recalibração, medida pelo teste de Hosmer-Lemeshow, é alcançada numa maior proporção, quando se utilizam diagnósticos principais, desagregados a 4 ou 5 dígitos (37.4%);

- Ao modelo M8, na qual a recalibração é atingida quando se considera o diagnóstico principal a 3 dígitos (18.7%);
- Aos modelos M 9 e M13, representando respectivamente, 9.8% e 9.4% do total de modelos de regressão logística apurados. Como principal particularidade, deve referir-se que estes modelos representam desagregações dos modelos M8 e M12, aos quais foram acrescentados as variáveis idade;
- A consideração do total de modelos que incluem o termo idade (modelos M3, 6, 7, 9, 11, 13 e 14) permite identificar 58 análises de regressão logística, as quais representam cerca de 27% do total apurado. Tal facto permite evidenciar que, pese embora o facto de a idade ser considerada no modelo inicial do Disease Staging, este factor parece ter, pelo menos para algumas doenças, efeitos diferentes nos E.U.A. e em Portugal;
- Em 17 situações, cerca de 8%, do total de modelos de regressão logística apurados, não foi possível alcançar calibração (medida pelo teste de Hosmer-Lemeshow), pelo que se realizará a uma análise mais detalhada das situações de saúde em que tal ocorre.

Quadro XI
Critérios para a Recalibração do Disease Staging
Casos Médicos

Critério de Regressão	Número de Ocorrências	% de Ocorrências
M1	0	0,00
M2	7	3,27
M3	2	0,93
M4	2	0,93
M5	4	1,87
M6	9	4,21
M7	2	0,93
M8	40	18,69
M9	21	9,81
M10	3	1,40
M11	4	1,87
M12	80	37,38
M13	20	9,35
M14	3	1,40
M15	17	7,94
Total	214	

No Quadro XII são identificadas e caracterizadas as situações em que não foi possível atingir calibração (teste de Hosmer-Lemeshow).

Quadro XII
Situações sem Calibração
Casos Médicos

Doença	Casos	TxMort Obs	TxMort Esp	Razão
Tumor do antro pilórico	858	23.54	12.20	1.93
Tumor de outros locais específicos do estômago	2358	35.88	14.35	2.50
Tumor do recto	3005	17.70	8.79	2.01
Tumor de outros locais específicos da laringe	733	18.69	10.29	1.82
Tumor do lobo superior, brônquios ou pulmão	1399	22.73	17.41	1.31
Tumor de outros locais dos brônquios ou pulmão	4430	34.88	16.92	2.06
Tumor dos brônquios e do pulmão, inespecífico	2149	36.11	16.80	2.15
Melanoma do membro inferior, incluindo a anca	107	33.64	11.95	2.82
Tumor de outros locais específicos da mama feminina	2576	34.67	12.56	2.76
Tumor do útero	5	40.00	13.50	2.96
Tumor de outros locais específicos da bexiga	1986	16.87	9.57	1.76
Tumor da bexiga, inespecífico	499	24.85	10.61	2.34
Tumor do rim, excepto pélvis	904	21.79	7.80	2.79
Doenças cerebrovasculares agudas, mal definidas	2262	39.35	9.48	4.15
Pneumonia bacteriana, inespecífica	13306	13.79	4.70	2.93
Broncopneumonia por microrganismo não especificado	10521	16.53	5.21	3.17
Pneumonia por microrganismo não especificado	37906	16.92	5.14	3.29
Total	85004	19.79	7.20	2.75

Das 17 situações constata-se que, na generalidade, todas correspondem a doenças com altas taxas de mortalidade, com um valor mínimo de 13.8%, um valor máximo de 40.0% e um valor médio de 19.8%.

A estas 17 situações correspondem 85004 episódios de internamento, os quais representam somente 3.7% do total de doentes envolvidos na análise e 6.0% dos doentes considerados como casos médicos.

Em relação à razão entre os valores observados e os previstos, verifica-se um valor mínimo de 1.31 e um valor máximo de 4.15, sendo o valor médio de 2.75. Estes valores, só por si, não podem ser considerados como o principal factor justificativo para a falta de calibração, visto que o valor médio para todos os doentes tratados é de 2.48.

Para o total de episódios de internamento em que não foi possível obter recalibração (85004 casos), 72.6% correspondem a doenças respiratórias (e todas elas pertencentes ao grupo da Pneumonia e gripe, de acordo com a CID-9-MC), 24.7% a Tumores malignos (ver Quadro XII) e os restantes 2.7% às Doenças cerebrovasculares agudas, mal definidas.

Quanto aos tumores, a situação é a seguinte:

-
- Os 21009 episódios de internamento em que não foi possível obter recalibração representam 26.1% do total dos casos com tumores malignos (80341);
 - A taxa de mortalidade para os 13 diagnósticos principais sem recalibração (28.3%) é ligeiramente superior à observada para a totalidade dos doentes com tumores malignos – 25.1%;
 - Para a taxa de mortalidade esperada, resultante da aplicação directa do Disease Staging, o panorama é praticamente idêntico, com valores esperados para o subgrupo sem recalibração de 13.3%% e de 11.8% para a globalidade dos doentes com tumores malignos;
 - Em relação aos Tumores malignos do estômago (código 151 da CID-9-MC) os casos tratados nos diagnósticos principais sem recalibração (Tumor do antro pilórico e Outros locais específicos do estômago) representam cerca de 53% do total de episódios de internamento com esta condição, apresentando ainda uma taxa de mortalidade observada superior à dos restantes diagnósticos, respectivamente de 32.6% e de 27.5.5%. O mesmo ocorre para as taxas de mortalidade previstas pelo Disease Staging para o subgrupo de doenças sem recalibração e para os restantes diagnósticos dos doentes com Tumor maligno do estômago, visto que as mesmas são respectivamente de 13.8% e de 12.8%;
 - Para os Tumores malignos do recto, da junção rectosigmóide e do ânus (código 154) o diagnóstico principal sem recalibração – Tumor maligno do recto (código 1541) representa cerca de 70% do total de episódios de internamento com esta situação, com uma taxa de mortalidade observada de 17.7% nos casos sem recalibração, enquanto que as restantes doenças deste grupo apresentam uma taxa de mortalidade observada de 24.3%. Os valores esperados, derivados da aplicação directa do Disease Staging são de 8.8% e de 9.7%, respectivamente para o Tumor maligno do recto e para as restantes doenças dentro deste grupo;
 - Em relação aos Tumores malignos da laringe (código 161) o diagnóstico principal sem recalibração (Tumor de outros locais específicos da laringe, código 1618) representa 41% do total de doentes com esta doença, sendo ainda a taxa de mortalidade observada ligeiramente superior para o diagnóstico 1618, 18.7% em oposição a 14.2% para os restantes diagnósticos principais. Praticamente o mesmo comportamento é constatado quando se comparam os valores previstos pelo Disease Staging, 10.3% para o Tumor de outros locais específicos da laringe e 7.3% para as restantes diagnósticos incluídos nos Tumores malignos da laringe;
 - Para os Tumores malignos da traqueia, dos brônquios e do pulmão, observam-se três situações (Tumor do lobo superior, brônquios ou pulmão, Tumor de outros locais dos brônquios ou pulmão e Tumor dos brônquios e do pulmão, inespecífico) em que não existe recalibração, representando 82.7% do total de doentes. As taxas de mortalidade são igualmente superiores no subgrupo de diagnósticos sem recalibração

- (33.1% em oposição a 24.5% para os restantes diagnósticos). Situação relativamente diferente é encontrada nos valores previstos, em que as taxas de mortalidade são respectivamente de 17.0% para o grupo de doenças sem recalibração e de 17.1% para o total de doentes;
- No Melanoma maligno da pele, encontra-se somente um diagnóstico principal (Melanoma do membro inferior, incluindo anca) em que não existe recalibração. Este diagnóstico representa 31.9% do total de doentes tratados e apresenta uma taxa de mortalidade inferior à da totalidade dos doentes, respectivamente de 33.6% e de 38.6%. Situação semelhante é encontrada nos valores previstos de mortalidade, com uma taxa superior no Melanoma do membro inferior, incluindo anca de 12.0% em oposição a 12.6% para os restantes doentes;
 - Em relação ao Tumor maligno da mama feminina, encontra-se somente um diagnóstico principal sem recalibração (Tumor de outros locais específicos da mama feminina), o qual representa 61.1% do total de episódios de internamento. A taxa de mortalidade observada é superior à do restante grupo de diagnósticos, respectivamente de 34.7% e de 27.3%. O mesmo acontece com os valores esperados, visto que as taxas de mortalidade são de 12.6% e de 11.0% para os diagnósticos sem e com recalibração;
 - Para o Tumor maligno da bexiga, são encontrados três diagnósticos principais (Tumor do útero, Tumor de outros locais específicos da bexiga e Tumor da bexiga, inespecífico) sem recalibração. Estes diagnósticos representam 87.6% do total de doentes e apresentam uma taxa de mortalidade superior à dos restantes diagnósticos, respectivamente de 18.5% e de 10.3%. Para a taxa de mortalidade esperada as diferenças entre os dois grupos são mais reduzidas, de 9.8% para o grupo de diagnósticos sem recalibração e de 9.3% para os restantes diagnósticos;
 - Para os Tumores Malignos do rim, de órgãos urinários e de órgãos urinários não especificados, somente o Tumor do rim, excepto pélvis não foi recalibrado, representando no entanto 85.9% do total de doentes tratados. A taxa de mortalidade observada é inferior no diagnóstico sem recalibração (21.8% e 25.0% nos restantes diagnósticos), enquanto que a taxa de mortalidade esperada é praticamente igual para os dois grupos (7.8% no diagnóstico sem recalibração e 7.9% nos restantes diagnósticos).

No que se refere às Doenças cerebrovasculares agudas, mal definidas (código 436), constitui o único diagnóstico principal nas Doenças circulatórias sem recalibração. Fazendo a comparação somente no que se refere às Doenças cerebrovasculares (códigos 430 a 438), observa-se que estes casos somente representam 2.7% do total de doentes com Doença cerebrovascular. Em relação às taxas de mortalidade observada constata-se uma grande diferença entre os dois subgrupos, com as Doenças cerebrovasculares agudas, mal definidas a apresentar valores de 39.3%, enquanto que para os restantes diagnósticos, a

taxa de mortalidade é de 16.3%. Embora esta tendência se mantenha para os valores previstos, respectivamente de 9.5% e de 6.1%, existe uma maior aproximação entre os dois grupos de diagnósticos.

Para as Doenças respiratórias somente se encontram problemas de recalibração em três situações (Pneumonia bacteriana não especificada, Broncopneumonia por microrganismo não especificado e Pneumonia por microrganismo não especificado). Atendendo a que todas estas doenças estão incluídas no Subgrupo de Doenças e Lesões “Pneumonia e gripe”, todas as comparações que se irão fazer respeitam somente a esta população. Estas três situações representam 69.3% daqueles episódios de internamento, apresentando ainda os diagnósticos sem recalibração taxas de mortalidade superiores (17.0%) às dos restantes diagnósticos (12.1%). Comportamento idêntico se observa para os valores previstos, respectivamente de 6.3% e de 2.1%, para os diagnósticos sem e com recalibração.

A análise do processo de recalibração dos casos médicos, permite afirmar que em 6% dos doentes não foi possível atingir este objectivo. Após uma análise mais detalhada sobre as situações em que tal ocorreu, não se conseguiu encontrar um padrão bem definido para se justificar esta falta de calibração dos dados do Disease Staging em Portugal.

Em termos de síntese, pode afirmar-se que, neste 1º Passo, o processo de recalibração para os casos médicos resultou mais complexo que nos casos cirúrgicos. Na realidade, enquanto que para os 897559 casos cirúrgicos foram apurados novos valores para a mortalidade prevista através de 88 modelos de regressão logística, para os casos médicos, envolvendo 1417630 episódios de internamento, tiveram de ser concretizados 214 modelos de regressão logística distintos, para se estimarem os novos valores para a mortalidade prevista.

Os subgrupos onde se tiveram de se apurar mais modelos de regressão logística respeitam aos Tumores (54.7%), às Doenças do aparelho circulatório (21.5%) e às Doenças do aparelho respiratório (13.1%), representando no seu conjunto cerca de 89% do total de diferentes modelos de regressão logística concretizados.

Assim, pode afirmar-se que para o total dos episódios de internamento se observaram 472502 (20.4%) em que não foi possível obter recalibração no 1º passo.

2º Passo

Em função dos novos valores previstos para a mortalidade derivados da metodologia do 1º Passo, procedeu-se a um novo reajustamento do modelo,

utilizando o novo valor do “logit p” e com tratamento separado para os casos cirúrgicos e médicos.

O “logit p” é calculado em função da mortalidade prevista da análise de regressão logística do 1º Passo.

Com este novo Passo pretende-se essencialmente averiguar se o ajustamento do modelo (calibração e discriminação) é melhorado, tanto para os casos cirúrgicos, como para os casos médicos, essencialmente nas situações em que não foi possível atingir calibração no 1º Passo.

Para os casos cirúrgicos seguiu-se a seguinte metodologia:

1. Regressão por Grandes Categorias de Diagnóstico (ver Anexo II) – C1
2. Se não existir calibração, introduzir variável DRG – C2
3. Se não existir calibração, introduzir variável idade – C3
4. Se não existir calibração, introduzir variáveis idade e DRG – C4
5. Se não existir calibração, introduzir variável PDXCAT – C5

Para os casos médicos seguiu-se a seguinte metodologia:

1. Regressão pelos Grandes Agrupamentos de Doenças (GADs) (Ver Anexo I) – M1
2. Se não existir calibração, introduzir variável PDXCAT – M2
3. Se não existir calibração, introduzir variável idade – M3
4. Se não existir calibração, introduzir variáveis PDXCAT e idade – M4
5. Se não existir calibração, proceder à regressão pelo PDXCAT – M5
6. Se não existir calibração, introduzir variável diagnóstico principal - M6
7. Se não existir calibração, introduzir variável idade – M7
8. Se não existir calibração, introduzir variáveis diagnóstico principal e idade – M8
9. Casos sem calibração – M9

Os resultados da aplicação da técnica de regressão logística para os casos cirúrgicos são apresentados no Quadro XIII, onde os dados são apresentados por Grandes Categorias Diagnósticas segundo a classificação dos DRGs (ver Anexo II, para identificação destas categorias).

A avaliação dos dados do Quadro XIII permite concluir que na grande maioria das situações o tipo de regressão logística mais frequente respeita à utilização do “logit p” e de um termo de DRG, como variáveis explicativas, o que aconteceu em 52% das situações (ver Quadro XIV).

Quadro XIII
Ajustamento do Modelo Disease Staging (2º passo)
Casos Cirúrgicos

Grandes Categorias Diagnósticas	Modelo	Calibração	Discriminação
GCD 0	C1	0.164	0.856
GCD 1	C2	0.231	0.937
GCD 2	C1	0.382	0.598
GCD 3	C2	0.999	0.914
GCD 4	C1	0.166	0.828
GCD 5	C2	0.140	0.897
GCD 6	C2	0.844	0.928
GCD 7	C2	0.106	0.943
GCD 8	C4	0.633	0.928
GCD 9	C2	0.470	0.973
GCD 10	C2	0.932	0.846
GCD 11	C2	0.900	0.853
GCD 12	C2	0.974	0.832
GCD 13	C2	0.984	0.940
GCD 14	C2	0.999	0.886
GCD 16	C1	0.699	0.859
GCD 17	C2	0.399	0.816
GCD 18	C1	0.493	0.897
GCD 19	C1	0.844	0.856
GCD 21	C2	0.075	0.844
GCD 22	C3	0.133	0.838
GCD 23	C1	0.404	0.857
GCD 24	C1	0.174	0.842
GCD 25	C3	0.464	0.561
GCD outros	C5	0.961	0.941

Quadro XIV
Critérios para a Recalibração do Disease Staging (2º passo)
Casos Cirúrgicos

Critério de Regressão	Número de Ocorrências	% de Ocorrências
C1	8	32
C2	13	52
C3	2	8
C4	1	4
C5	1	4
Total	25	100

A calibração, medida pelo Teste de Hosmer Lemeshow, foi atingida nas 25 regressões logísticas, sendo ainda de referir que em 9 situações a significância deste teste foi superior a 0.750. Na situação contrária, encontra-se uma situação com valor inferior a 0.099 e 6 regressões com valores entre 0.100 e 0.200.

O valor com calibração mais baixa, embora com valores significativos, ocorreu na Grande Categoria Diagnóstica 21 – Traumatismos, intoxicações e efeitos tóxicos da droga, na qual estão considerados 5675 episódios de internamento com uma taxa de mortalidade de 1.69%.

As restantes 6 situações com valores do Teste de Hosmer-Lemeshow com níveis de significância mais baixos são encontrados nas Grandes Categorias Diagnósticas 0, 4, 5, 7, 22 e 24, são responsáveis por 120473 episódios de internamento, 13.6% do total de casos cirúrgicos, dentro dos quais as Doenças do Aparelho Circulatório são as mais frequentes (67955 doentes tratados). Por outro lado, as taxas de mortalidade são muito elevadas (com valores superiores a 20%) em três Grupos de Doenças – Pré-grandes categorias diagnósticas, Queimaduras e Traumatismos múltiplos significativos.

Com excepção de duas situações, todos os ajustamentos apresentam um bom poder discriminativo, medido pela estatística “c”. Nas Doenças e perturbações do olho, embora tenham sido tratados 70472 doentes a taxa de mortalidade é muito baixa – 0.03%. Situação perfeitamente oposta é encontrada nas Infecções pelo Vírus da imunodeficiência humana, com poucos doentes tratados (171 episódios), embora com uma taxa de mortalidade elevada – 25.7%.

Assim, pode concluir-se que, para os casos cirúrgicos, embora se devam ter presentes as considerações anteriormente referidas, a recalibração do Disease Staging para os dados portugueses não suscitou grandes problemas, pelo que se podem comparar os valores previstos e observados para a mortalidade e desta forma avaliar a efectividade dos hospitais.

Os resultados da aplicação da regressão logística para os casos médicos são apresentados no Quadro XV.

Para as 34 análises de regressão logística, os modelos mais frequentes respeitam à utilização do “logit p” e PDXCAT, como variáveis explicativas (10 situações) e do “logit p”, como variável explicativa (8 situações).

Deve ainda evidenciar-se que somente em quatro situações não foi atingida recalibração. Pela sua importância estes aspectos serão analisados separadamente.

Praticamente em todas as situações os modelos apresentam um bom valor discriminativo, sendo de referir que os valores mais baixos são encontrados na Infecção por VIH, Parto normal e Admissão para quimioterapia.

Quadro XV
Ajustamento do Modelo Disease Staging (2º passo)
Casos Médicos

Doenças	Modelo	Calibração	Discriminação
Tumor maligno da mama feminina	M9	-----	-----
Doenças da mama (restantes)	M1	0.888	0.982
Doença cerebrovascular	M9	-----	-----
Doenças do sistema nervoso central (restantes)	M1	0.110	0.912
Doenças cardiovasculares e do coração	M2	0.760	0.838
Doenças dermatológicas	M1	0.719	0.929
Doenças do ouvido, nariz e garganta	M1	0.587	0.927
Doenças gastrointestinais	M3	0.054	0.913
Doenças ginecológicas	M2	0.584	0.906
Doenças hepatobiliares	M3	0.146	0.886
Mieloma múltiplo	M5	0.090	0.729
Doenças hematológicas (restantes)	M1	0.088	0.861
Imunodeficiências	M6	0.989	0.970
Infecção por VIH (vírus da imunodeficiência humana)	M5	0.820	0.676
Doenças linfáticas	M2	0.483	0.739
Doenças dos órgãos genitais masculinos	M4	0.847	0.914
Doenças musculoesqueléticas	M2	0.445	0.934
Doenças endócrinas e metabólicas	M1	0.120	0.819
Parto normal	M5	0.328	0.653
Doenças obstétricas (restantes)	M2	0.825	0.868
Doenças oftalmológicas	M1	0.678	0.907
Doenças do sistema nervoso periférico	M2	0.609	0.920
Doenças psiquiátricas	M2	0.412	0.867
Doenças renais	M3	0.297	0.905
Tumor maligno do pulmão, brônquios e mediastino	M8	0.159	0.690
Pneumonia bacteriana, outras	M9	-----	-----
Doenças respiratórias (restantes)	M4	0.204	0.829
Doenças multisistémicas	M4	0.780	0.898
Doenças vasculares (excepto coração)	M2	0.636	0.905
Doenças não especificadas	M2	0.790	0.838
Admissão para quimioterapia	M7	0.279	0.645
Tumor maligno: outros locais, não anteriormente classificados	M8	0.940	0.848
Outras doenças (restantes)	M4	0.609	0.948

Embora todos os ajustamentos apresentem calibração deve referir-se que em 3 situações, o seu valor significativo é inferior a 0.100 e em outras 3, se situa entre 0.100 e 0.200.

Para o Tumor maligno da mama feminina, Doença cerebrovascular e Pneumonia bacteriana (outras) não foi possível atingir a recalibração, pelo que se procedeu

a uma análise mais desagregada. Esta consistiu essencialmente na identificação dos diagnósticos principais que apresentavam um padrão diferente em termos de taxa de mortalidade e proceder a regressões logísticas separadas.

Para o Tumor maligno da mama feminina foram criados dois sub-grupos de Doenças, o diagnóstico principal (código 174) e os restantes diagnósticos principais. Os resultados deste novo procedimento são apresentados no Quadro XVI.

Quadro XVI
Ajustamento do Modelo Disease Staging
Tumor maligno da mama feminina, Casos Médicos

Doenças	Calibração	Discriminação
Tumor de outros locais específicos da mama feminina	0.000	0.698
Tumor da mama feminina, inespecífico	0.154	0.652
Tumor da mama feminina (restantes)	0.193	0.860
Restantes	0.134	0.742

Para as cinco situações consideradas, somente não foi possível encontrar calibração no Tumor de outros locais específicos da mama feminina. A discriminação também não é muito forte no Tumor de outros locais específicos da mama feminina e no Tumor da mama feminina (inespecífico).

Para as Doenças cerebrovasculares, observem-se os valores do Quadro XVII.

Quadro XVII
Ajustamento do Modelo Disease Staging
Doença Cerebrovascular, Casos Médicos

Doenças	Calibração	Discriminação
Hemorragia subaracnoideia	0.062	0.763
Hemorragias intracranianas	0.127	0.748
Outras hemorragias intracranianas e as não especificadas	0.284	0.635
Oclusão e estenose das artérias pré-cerebrais	0.175	0.831
Oclusão das artérias cerebrais	0.066	0.708
Isquemia cerebral transitória	0.251	0.809
Doenças cerebrovasculares agudas, mal definidas	0.001	0.726
Outras doenças cerebrovasculares e as mal definidas	0.063	0.696
Efeitos tardios de doenças cerebrovasculares	0.117	0.784
Restantes	0.110	0.912

Somente nas Doenças cerebrovasculares agudas, mal definidas (código 436) não foi possível alcançar calibração. As doenças Outras hemorragias intracranianas e as não especificadas (código 432) e Outras doenças

cerebrovasculares e as mal definidas (código 437) apresentam valores de discriminação não muito fortes.

Para a “Pneumonia bacteriana, outras” foram basicamente identificadas três situações:

- Broncopneumonia por microrganismo não especificado (código 485);
- Pneumonia por microrganismo não especificado (código 486);
- Restantes Diagnósticos Principais incluídos nesta Categoria Principal de Doença.

Os resultados deste novo procedimento são apresentados no Quadro XVIII.

Quadro XVIII
Ajustamento do Modelo Disease Staging
Pneumonia bacteriana, outra, Casos Médicos

Doenças	Calibração	Discriminação
Broncopneumonia por microrganismo não especificado	0,810	0.802
Pneumonia por microrganismo não especificado	0.011	0.786
Restantes	0.073	0.787

Nesta Categoria Principal de Doença foi possível encontrar calibração no Diagnóstico Principal – Broncopneumonia por microrganismo não especificado (código 485). Para o subgrupo “restantes diagnósticos”, embora exista calibração, o valor do Teste de Hosmer-Lemeshow é baixo.

Para a Pneumonia por microrganismo não especificado não foi possível atingir calibração (código 486), pelo que após este 2º Passo, somente 37906 episódios de internamento não estão calibrados.

Para qualquer destas três situações o poder discriminativo é bom, com a estatística “c” a apresentar valores iguais ou superiores a 0.786.

Atentos os valores atrás referidos pode concluir-se que o processo de recalibração do Disease Staging aos dados portugueses foi concluído com sucesso, embora se deva salientar que se encontraram mais dificuldades nos casos não cirúrgicos.

Deve ainda ser referido que a consideração deste 2º Passo demonstrou grande utilidade, visto que para a grande maioria das situações sem calibração encontradas no 1º Passo, tanto para os casos cirúrgicos, como para os médicos, foi possível atingir recalibração. As exceções são encontradas no Tumor da mama feminina, outros locais específicos, nas Doenças cerebrovasculares agudas, mal definidas e na Pneumonia por microrganismo não especificado, o

que não permite inviabilizar a utilização do Disease Staging como metodologia para se apurar a Taxa de Mortalidade Esperada por hospital em Portugal.

A opção pelos resultados do 2º Passo da regressão logística é ainda justificada pela dimensão dos episódios de internamento/doentes sem recalibração. De facto, no 1º Passo, estavam nesta situação 472502 episódios de internamento, dos quais 85004 correspondem a casos médicos, enquanto que no 2º Passo somente existem somente 42744 episódios onde não foi possível atingir a recalibração, todos correspondentes a casos médicos.

Estes casos correspondem a 1.8% do total dos episódios de internamento e a 3% dos casos médicos. Pode ainda referir-se que entre os 1º e 2º passos para a recalibração, se observou uma diminuição em cerca de 90% dos episódios sem recalibração e de cerca de 50% quando se consideram os casos médicos.

Assim, a avaliação da efectividade dos hospitais portugueses, em função dos resultados de saúde e medidos pela comparação entre taxas de mortalidade observadas e esperadas vai ser realizado com recurso aos elementos do Disease Staging, os quais derivam de dados de rotina existentes em Portugal.

Recalibração da Duração de Internamento Esperada

Outro produto do “software” do Disease Staging é o estabelecimento de previsões para a duração de internamento por doente.

Segundo os seus autores o modelo empírico deriva da utilização de uma base de dados de cerca de 14 milhões de doentes e com a seguinte equação (MEDSTAT, 2001):

Duração de Internamento Esperada = Peso da Comorbilidade + Peso da Comorbilidade ao Quadrado + Raiz Quadrada do Peso da Comorbilidade + Idade + Raiz Quadrada da Idade + Idade ao Quadrado + Idade Igual a Zero + Sexo + Admissões Electivas + Admissões Urgentes + Transferências + Número de Diagnósticos + Sub-estadio do Diagnóstico Principal.

NOTA: Idade = 0 é uma variável “dummy” para distinguir os recém nascidos dos outros doentes, assumindo o valor 1 quando a idade é igual a zero e o valor 0 quando a idade >0.

O Peso da Comorbilidade por sua vez é função do diagnóstico e do respectivo sub-estadio, do sexo, da idade, da raiz quadrada da idade, da idade ao quadrado e da idade igual a zero.

NOTA: Este Peso da Comorbilidade é diferente do calculado para a mortalidade, visto que o efeito das comorbilidades é diferente na duração de internamento.

Segundo estes autores (MEDSTAT, 2001) o Peso da Comorbilidade para a Duração de Internamento foi baseado em regressões para cada DRG no logaritmo (Ln) das “charges”, visto as mesmas estarem significativamente correlacionadas com a duração de internamento (as “charges” correspondem à facturação de cada doente).

Assim, o Peso da Comorbilidade para a Duração de Internamento é igual

$$\text{Ln} \left(1 + \frac{\sum_{j=1}^k C_j}{\text{CPDXCAT}} \right)$$

Em que Ln é o logaritmo natural, C_j é a “charge” prevista para a comorbilidade J e CPDXCAT é a “charge” prevista para a doença principal (as precisões são ajustadas pelo estadió da doença, idade e sexo). Neste sentido, o quociente dentro de parênteses é igual à soma das “charges” médias para cada comorbilidade a dividir pelas “charges” totais da doença principal. Quando, um doente não tem comorbilidades tem um Peso de Comorbilidade de $\text{Ln}(1) = 0$ (MEDSTAT, 2001).

Ainda segundo os autores este modelo de regressão linear conduziu a um coeficiente de determinação da duração de internamento de 0.36 (MEDSTAT, 2001).

Uma aplicação directa dos resultados desta metodologia do Disease Staging aos hospitais portugueses é apresentada no Quadro XIX.

Mais uma vez o problema de calibração está presente, com uma razão entre valores observados (demora média real dos hospitais portugueses) e os valores esperados (demora média esperada em função dos dados originais do modelo) de 2.2.

Este comportamento é relativamente constante entre os Grandes Agrupamentos de Doenças, com um coeficiente de variação de 0.16, embora com um valor máximo de 3.14 para as doenças do sistema imunológico e com um mínimo de 1.55 para os casos obstétricos.

Esta disparidade entre os valores observados e esperados parece estar mais associada ao comportamento dos hospitais portugueses – o coeficiente de correlação entre a demora média e a razão entre valores observados e esperados é de 0.643 – do que aos valores de origem do Disease Staging, o coeficiente de correlação entre valores esperados e a razão é de 0.336.

Quadro XIX
Demora Média Observada e
Demora Média Esperada por Doença
e para o Total de Doentes/Episódios de Internamento

Grandes Agrupamentos de Doenças	Demora Média Observada	Demora Média Esperada	Razão
Doenças da mama	6.68	2.47	2.71
Doenças do sistema nervoso central	9.73	3.99	2.44
Doenças cardiovasculares e do coração	7.53	3.36	2.24
Doenças dermatológicas	6.41	3.14	2.04
Doenças do ouvido, nariz e garganta	4.63	2.14	2.17
Doenças gastrointestinais	6.78	3.32	2.05
Doenças ginecológicas	5.30	2.47	2.14
Doenças hepatobiliares	8.66	3.71	2.33
Doenças hematológicas	10.89	4.92	2.21
Doenças Imunológicas	16.70	5.32	3.14
Doenças linfáticas	13.28	5.49	2.42
Doenças dos órgãos genitais masculinos	6.11	2.61	2.34
Doenças musculoesqueléticas	9.31	3.19	2.91
Doenças endócrinas e metabólicas	9.07	3.89	2.33
Doenças obstétricas	3.55	2.29	1.55
Doenças oftalmológicas	3.40	1.53	2.21
Doenças do sistema nervoso periférico	3.87	2.06	1.88
Doenças psiquiátricas	14.53	6.25	2.32
Doenças renais	8.26	3.39	2.43
Doenças respiratórias	10.21	4.72	2.16
Doenças multisistémicas	8.25	3.29	2.51
Doenças vasculares (excepto coração)	8.12	3.29	2.47
Doenças não especificadas	6.30	3.76	1.68
Outras doenças	4.59	2.79	1.64
Total	7.18	3.26	2.20

Perante esta situação duas alternativas se colocam:

- Recalibrar os valores esperados para a demora média à realidade portuguesa mediante a realização de regressões lineares;
- Recalibrar os valores esperados do Disease Staging para a demora média, ponderados por um valor de referência.

Para a recalibração do Disease Staging à realidade portuguesa com recurso à regressão linear, uma estratégia possível é a seguinte (MEDSTAT, 2002):

- Transformar a duração de internamento (DI) = 0 em DI = 1;
- Eliminar “outliers” por DRG;

- Fazer Regressões Lineares por DRG com a seguinte equação:

$$\text{Ln (DIO)} = a + b * \text{Ln (DIE)}$$
 Em que:
 Ln – Logaritmo Natural
 DIO – Duração de Internamento Observada
 DIE – Duração de Internamento Esperada
- Estimar a Nova Duração de Internamento (NDIE) de acordo com a seguinte equação:

$$\text{NDIE} = f * \exp (a + b * (\text{Ln DIE}))$$
 Em que:
 f é um factor de transformação e de correcção, designado por estimativa “smearing”, calculado da seguinte forma: f = média das exponenciais dos resíduos da equação anterior
 a e b são os coeficientes de regressão da equação anterior.

Para a segunda alternativa, recalibrar os valores estimados para a duração de internamento com recurso a um valor de referência, uma estratégia possível é a seguinte (MEDSTAT, 2002):

- Transformar a duração de internamento (DI) = 0 em DI = 1;
- Eliminar “outliers” por DRG;
- Calcular a demora média observada por DRG (excluindo “outliers”)

$$A = \text{Dias de Internamento Observados} / \text{Doentes Saídos};$$
- Calcular a demora média esperada por DRG (excluindo “outliers”)

$$B = \text{Dias de Internamento Esperados} / \text{Doentes Saídos};$$
- Calcular a razão entre valores observados e esperados para a demora média por DRG (excluindo os “outliers”)

$$C = A / B;$$
- Multiplicar a duração de internamento esperada (dados originais do Disease Staging) pelos valores encontrados em C e desta forma estimar a nova duração de internamento esperada (excluindo “outliers”).

Neste estudo opta-se pela segunda alternativa, ou seja recalibrar a nova duração de internamento em função de um valor de referência. Esta estratégia não conduz a enviesamento no apuramento dos resultados, visto que, se utiliza sempre o mesmo procedimento para todas as perspectivas de análise e para todos os hospitais.

A identificação de “outliers” é feita por DRG de acordo com a seguinte equação (Emerson e Strenio, 1983):

“Outliers” inferiores = valores inferiores a $FL - 1.5 * AIQ$

“Outliers” superiores = valores superiores a $FU + 1.5 * AIQ$

Em que: FL – 1º Quartil

FU – 3º Quartil

AIQ – Amplitude Interquartil (FU – FL)

Para se identificarem todos os “outliers” procedeu-se em primeiro lugar a uma transformação de todos os dias de internamento igual a 0 em 1, mantendo-se as restantes durações de internamento com valores iguais aos das bases de dados.

Atendendo à dimensão da população em estudo – 489 DRGs – no Anexo III são apresentados por DRG, o total de doentes saídos, a média, a mediana, o 1º e 3º quartis, a distância quartal, e as barreiras de “outliers” inferiores e superiores.

Neste sentido, o resumo destes procedimentos inclui somente a enumeração das barreiras de “outliers” inferiores e superiores, bem como do número de “outliers” por DRG.

Em relação às barreiras inferiores de “outliers” deve, em primeiro lugar referir-se que em 472 dos 489 DRGs (96.5%) estes valores são negativos e que em 2 DRGs (nºs 372 e 378) são iguais a zero. Nos DRGs 41, 59, 105, 163, 232, 270, 323, 343, 351 e 481 a barreira é igual a um, portanto sem “outliers” inferiores. Por sua vez, no DRG 337 a barreira é igual a dois, nos DRGs 109 e 371 igual a três, no DRG 436 igual a quatro.

Para as barreiras de “outliers” superiores existem 4 DRGs (41, 163, 270 e 343) com valores iguais a 1, correspondendo ao DRG 103 o valor mais elevado (201 dias).

No Anexo III são apresentados os “outliers”, bem como as respectivas percentagens por DRG.

Em termos globais esta metodologia identificou 138973 casos que são “outliers” em relação à duração de internamento. Nos DRGs 41, 163, 343, 36, 129, 270, 125, 492, 156 e 409 foram encontradas as maiores percentagens de “outliers”, com 24.9% no primeiro e 13.8% no último. Por sua vez os DRGs 340, 70, 126, 322, 56, 309, 291, 505, 314 correspondem às situações com menos “outliers”, com 0.02% no primeiro e 1.8% no último.

No Anexo III são apresentadas as demoras médias por DRG para o total dos doentes e para a população sem “outliers”. No mesmo Anexo, pode igualmente ser consultada a demora média esperada para a população sem “outliers”.

Esta demora média esperada corresponde aos valores originais do Disease Staging e consequentemente está calibrada para os dados dos EUA. A metodologia que se adoptou para se recalibrarem os dados à realidade portuguesa (ver a equação referida anteriormente) implica que por DRG a demora média esperada é igual à observada, variando somente os dados por caso, em função das características dos doentes.

Tendo em atenção os dados apresentados no Anexo III pode referir-se que nos DRGs 129, 156, 465, 352, 431, 299, 256, 467, 9 e 409 foi onde se observou uma maior redução na demora média quando se compara o total de doentes saídos

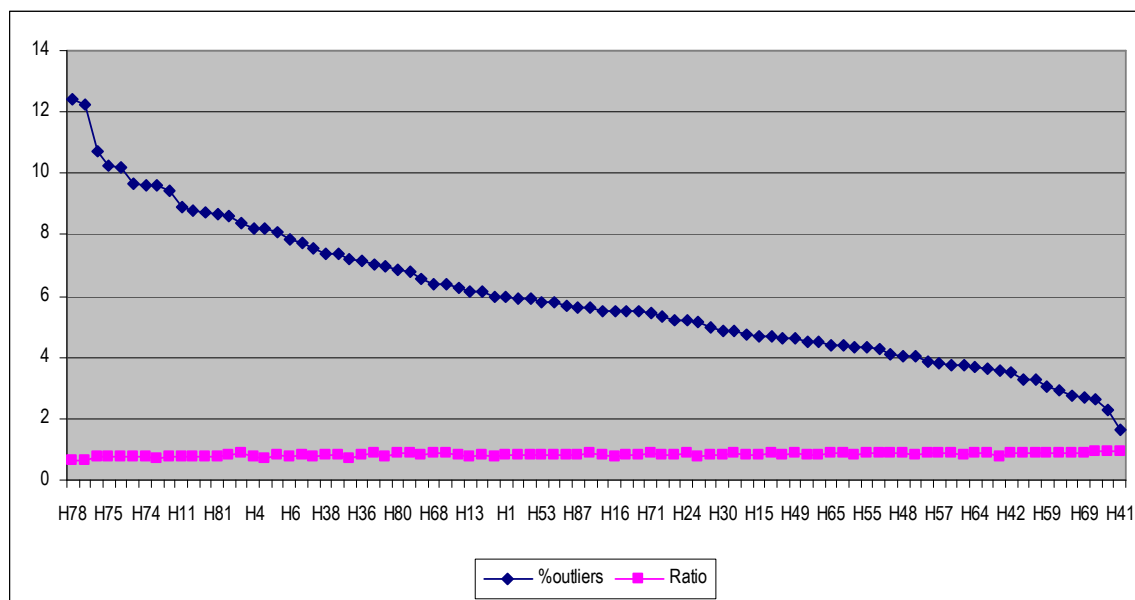
com a população excluindo os “outliers”. A situação contrária, ou seja, menor redução da demora média com e sem “outliers” foi encontrada nos DRGs 340, 126, 70, 56, 106, 322, 167, 291, 237 e 309.

A comparação entre a demora média observada e a demora média esperada por DRG para todos os doentes excluindo os “outliers” permite as seguintes observações:

- Nos DRGs 36, 471, 237, 344, 306, 307, 257, 217, 336 e 310 são encontradas as maiores diferenças entre estes indicadores, com valores superiores para a demora média observada;
- Nos DRGs 431, 351, 343, 270, 465, 129, 299, 467, 436 e 137 acontece a mesma situação, mas agora com valores superiores para a demora média esperada;
- Nos DRGs 125, 187, 360, 411, 118, 342, 352, 410, 462 e 184 são encontrados os valores mais aproximados entre a demora média observada e a demora média esperada em função dos dados originais do Disease Staging.

Em seguida são apresentadas, por hospital, as implicações da eliminação dos “outliers”, designadamente a percentagem de “outliers” por hospital e as reduções na demora média observada, igualmente por hospital (ver Figura 7). No Anexo III podem ser consultados os respectivos valores.

Figura 7
% de *Outliers* por Hospital; Reduções na Demora Média por Hospital

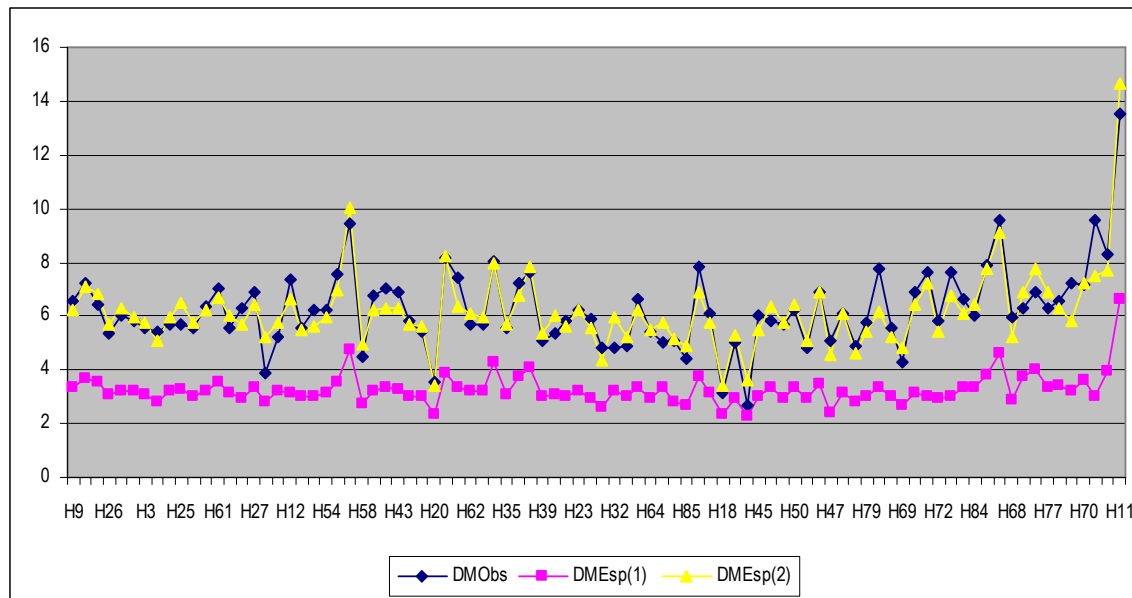


Como primeira observação deve referir-se que a variabilidade na percentagem de “outliers” é superior à observada nas reduções da demora média, sendo os coeficientes de variação de 0.38 e de 0.07, respectivamente. No entanto, apesar desta situação a correlação existente entre estas duas variáveis é forte, sendo o coeficiente de correlação de -0.81 (significativo a 0.01), o que exprime que quanto mais “outliers” maior é a redução na demora média por hospital.

Por outro lado, conforme foi referido a percentagem de “outliers” para o total dos hospitais é de 6.00%, com um máximo de 12.4% e um mínimo de (1.6%).

Finalmente, em função da metodologia descrita, calculou-se a demora média esperada por hospital, recalibrada aos dados portugueses. Na Figura 8 são apresentadas a demora média observada por hospital, a demora média esperada por hospital em função dos dados originais do Disease Staging (DMEsp(1)) e a demora média por hospital após recalibração aos dados portugueses (DMEsp(2)), estando os hospitais ordenados por decrescente da produção. No Anexo III podem ser consultados os respectivos valores.

Figura 8
Demora Média Observada por Hospital
Demora Média Esperada por Hospital (dados originais) (1)
Demora Média Esperada por Hospital (após recalibração) (2)



Como primeira constatação pode referir-se que a demora média observada por hospital está fortemente associada com a demora média prevista pelo Disease Staging, visto que o coeficiente de correlação é de 0.85 (significativo a 0.01).

Após a recalibração dos dados à realidade portuguesa observa-se igualmente uma forte associação com os dados originais (coeficiente de correlação de 0.96, igualmente significativo a 0.01).

Neste sentido, não é de estranhar que a demora média observada e a demora média prevista após a recalibração apresente igualmente um coeficiente de correlação elevado (0.92, igualmente significativo a 0.01).

No entanto, apesar destas associações fortes convém destacar o seguinte para a comparação entre a demora média observada e a demora média esperada com a aplicação directa do Disease Staging:

- Para a globalidade dos hospitais a demora média observada é 1.89 vezes superior à esperada;
- A variabilidade entre os hospitais não é muito grande visto que o coeficiente de variação entre esta razão é de 0.14, embora este ratio apresente um máximo de 3.19 e um mínimo de 1.20.

A comparação entre a demora média esperada quando se aplicam os valores originais e a demora média esperada após a recalibração aos dados portugueses permite as seguintes observações:

- Embora, como é natural, a demora média esperada após recalibração seja 1.89 vezes superior à esperada em função dos dados originais, a variabilidade entre os hospitais é ainda mais reduzida, visto que o coeficiente de variação deste ratio é de 0.08. O valor máximo deste ratio é de (2.49) e o mínimo de 1.45.

Estes valores, para além de expressarem as diferenças detectadas entre DRGs, permitem concluir que, apesar da forte associação entre os valores esperados para a demora média decorrentes dos dados originais e os esperados após a recalibração aos dados portugueses, esta última metodologia permite uma avaliação directa da eficiência dos hospitais portugueses, pelo que será o indicador a ser utilizado para esta finalidade.

RESULTADOS

- Produção Hospitalar
- Desempenho
- Síntese de Resultados

5. Resultados

5.1. Produção Hospitalar

5.1.1. Medição da Produção Hospitalar

O Índice de Casemix pretende ser um indicador que sintetiza a diversidade dos casos tratados em cada hospital (Lichtig, 1986).

Anteriormente foram enunciadas as questões teóricas, bem como as metodologias para se proceder ao cálculo do Índice de Casemix, através dos dois sistemas de classificação de doentes utilizados neste estudo – os Diagnosis Related Groups (DRGs) e o Disease Staging (DS).

Para se estudar se os dois sistemas de classificação de doentes estão a dar a mesma informação serão analisados os seguintes aspectos:

- Comparação dos Índices de Casemix por Tipo de hospital e por Região de Saúde;
- Comparação dos Índices de Casemix por Hospital;
- Comparação dos Índices de Casemix por Doenças;
- Comparação dos Índices de Casemix por Gravidade;
- Comparação dos Índices de Casemix por Idade.

No Quadro XX são apresentados os Índices de Casemix por Tipo de hospital e por Região de Saúde.

Quadro XX
Índices de Casemix segundo os DRGs e o Disease Staging
por Tipo de Hospital e por Região de Saúde

	Tipo de Hospital		Região de Saúde	
	DRGs	DS	DRGs	DS
I	1.33	1.12	1.01	1.15
II	0.81	0.07	1.00	1.08
III	0.98	0.95	1.08	0.96
IV	0.94	1.17	1.21	1.13
V	1.86	1.31	1.08	0.87
Total	1.12	1.00	1.12	1.00

Conforme se verifica, tanto por Tipo de hospital, como por Região de Saúde, os dois sistemas de classificação de doentes estão a dar indicações diferentes sobre a produção dos hospitais.

Por Tipo de hospital, embora exista concordância no que se refere aos hospitais com Índices de Casemix mais e menos elevados, observam-se algumas diferenças nas posições intermédias. A título exemplificativo, observe-se o que ocorre nos hospitais do Tipo IV, visto que têm o 4º Índice de Casemix mais elevado segundo os DRGs e o 2º valor mais elevado quando se utiliza o Disease Staging.

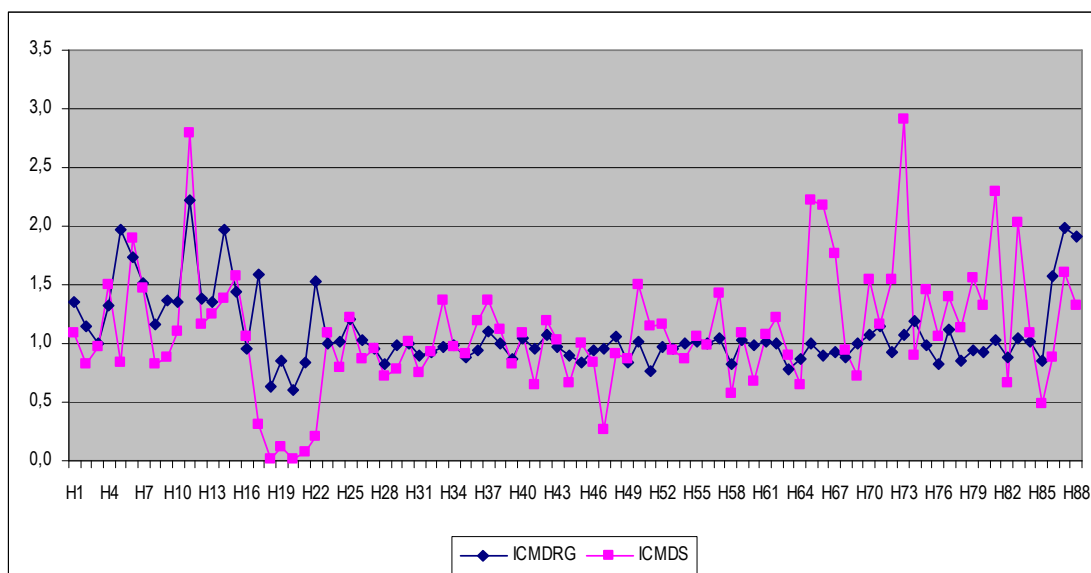
A este respeito deve ainda referir-se que o coeficiente de variação do Índice de Casemix entre os diversos Tipos de hospital é de 0.35, segundo os Diagnosis Related Groups e de 0.54 quando se utiliza o Disease Staging, o que parece evidenciar que existem diferenças mais significativas quando se utiliza este último sistema de classificação.

Por Região de Saúde, as diferenças são ainda mais relevantes, visto que existe uma discrepância importante na identificação das Regiões de Saúde com doentes com maior Índice de Casemix. De facto, enquanto os DRGs identificam por ordem decrescente as Regiões IV, III e V, como aquelas com Índices mais elevados, com o Disease Staging esta situação ocorre nas Regiões I, IV e II.

Contudo, a intensidade na variação do Índice de Casemix é menor do que a observada por Tipo de Hospital, sendo os coeficientes de variação de 0.08 e de 0.12, respectivamente para os Diagnosis Related Groups e para o Disease Staging.

Na Figura 9 são apresentados os Índices de Casemix por hospital, igualmente calculados em função dos DRGs e do Disease Staging. No Anexo IV podem ser consultados os respectivos valores.

Figura 9
Índices de Casemix segundo os DRGs e o Disease Staging por Hospital



Comparando os dois sistemas de classificação de doentes, observa-se que aparentemente estão a disponibilizar informação diferente, atribuindo os Diagnosis Related Groups Índices de Casemix superiores a 44 hospitais e o Disease Staging aos restantes 44 hospitais.

No entanto, atendendo a que o coeficiente de correlação entre os dois Índices de Casemix é de 0.34 (significativo a 0.01), pode concluir-se que por hospital, quanto maior é o Índice de Casemix calculado pelos DRGs, maior é este indicador quando é apurado pelo Disease Staging.

Quando a produção é avaliada pelo Índice de Casemix dos DRGs os hospitais apresentam um comportamento mais homogêneo do que quando se utiliza o Disease Staging, visto que os coeficientes de variação são de 0.28 e de 0.48, respectivamente.

Para as situações em que o Índice de Casemix dos DRGs é superior constata-se que, em termos relativos, são os hospitais do Tipo II, V e I, aqueles que apresentam maiores frequências.

Utilizando o Disease Staging são os hospitais do Tipo IV e III os que apresentam maiores frequências relativas.

Considerando agora os hospitais com Índice de Casemix mais elevado observa-se o seguinte:

- Com os DRGs são os hospitais 11, 87, 5, 14 e 88 os que apresentam valores mais elevados. No Disease Staging tal acontece nos hospitais 73, 11, 81, 65 e 66;
- Nos 10 hospitais com Índice de Casemix mais elevado, 5 são do Tipo I, 3 do Tipo V e 2 do Tipo II utilizando os DRGs. Com o Disease Staging, encontram-se 6 hospitais do Tipo IV, 3 do Tipo I e 1 do Tipo V;
- Para este sub-conjunto de hospitais os dois sistemas de classificação de doentes somente concordam em três. Ou seja, existem três hospitais identificados simultaneamente com Índices de Casemix mais elevados (hospitais 6, 11 e 87).

Para os hospitais com Índice de Casemix mais baixo a situação é a seguinte:

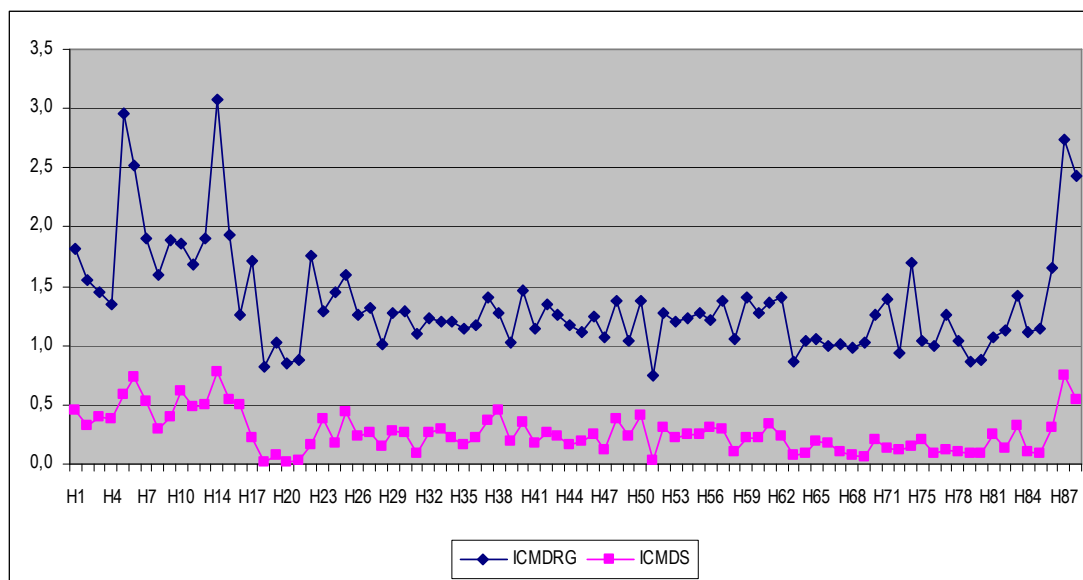
- Com o Disease Staging são os hospitais 18, 20, 21, 19 e 22 que apresentam valores mais baixos, enquanto que os DRGs identificam os hospitais 20, 18, 51, 63 e 28;
- Nos 10 hospitais com Índice de Casemix mais baixo, os DRGs identificam 5 do Tipo III, 3 do II e 2 do IV. O Disease Staging identifica 6 do Tipo II (todos os hospitais), 3 do Tipo III e 1 do IV;

- Existem 4 hospitais que são simultaneamente identificados com Índices de Casemix mais baixos pelos dois sistemas de classificação de doentes (hospitais, 18, 20, 21 e 58).

Os hospitais 22, 17, 5, 21 e 19 (4 do Tipo II e o restante do Tipo I) são os que apresentam maiores diferenças nos Índices de Casemix para valores superiores dos DRGs e os hospitais 73, 66, 81, 65 e 83 (todos do Tipo IV) para valores superiores do Disease Staging.

Na Figura 10 são apresentados os Índices de Casemix por hospital para os casos cirúrgicos, igualmente calculados em função dos DRGs e do Disease Staging. No Anexo IV podem ser consultados os respectivos valores.

Figura 10
Índices de Casemix segundo os DRGs e o Disease Staging por Hospital
Casos Cirúrgicos



Para os casos cirúrgicos a informação disponibilizada pelos dois sistemas de classificação de doentes é completamente diferente, visto que os DRGs atribuem sempre Índices de Casemix superiores aos calculados pelo Disease Staging.

Os hospitais 14, 5, 87, 6 e 88 são os que apresentam Índices de Casemix medidos pelos DRGs mais elevados. Com o Disease Staging estes valores são encontrados nos hospitais 14, 87, 6, 10 e 5. Estes dois sistemas de classificação concordam em 3 hospitais, valor que sobe para 8 hospitais quando se consideram os 10 hospitais com maior Índice de Casemix.

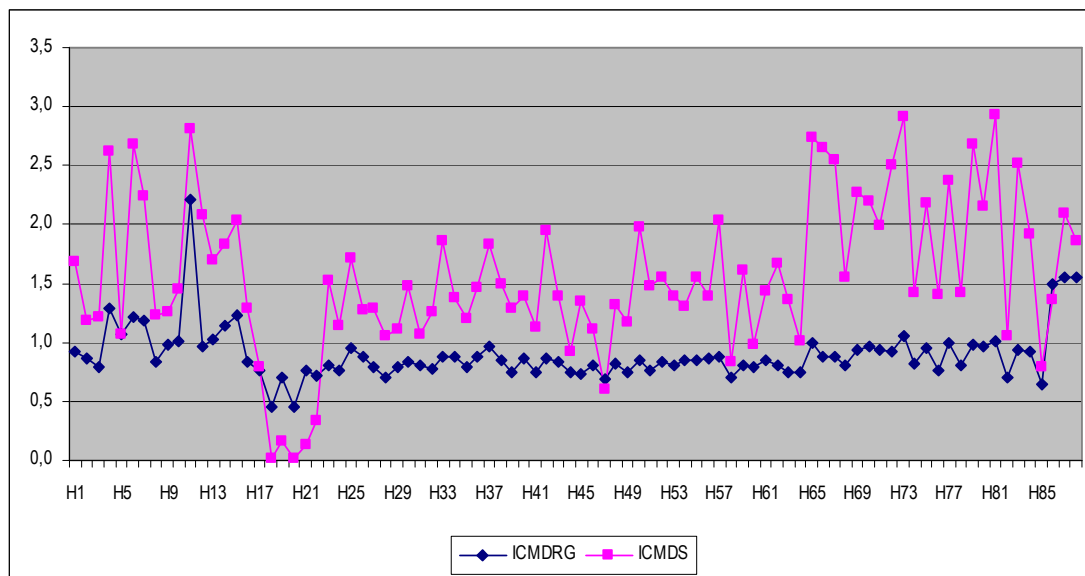
Nos 10 hospitais onde se encontram maiores diferenças nos Índices de Casemix medidos pelos DRGs e pelo Disease Staging, 5 são do Tipo I, 2 dos Tipos V e II e 1 do Tipo IV.

Quando a produção relativa aos casos cirúrgicos é avaliada pelo Índice de Casemix dos DRGs os hospitais apresentam um comportamento menos heterogêneo do que quando se utiliza o Disease Staging, visto que os coeficientes de variação são de 0.33 e de 0.65, respectivamente.

Contudo, apesar da situação descrita existe uma associação mais forte entre os Índices de Casemix calculados pelos DRGs e pelo Disease Staging do que o observado para o total dos doentes, visto que o coeficiente de correlação é de 0.851 (significativo a 0.01).

Na Figura 11 são apresentados os Índices de Casemix por hospital para os casos médicos, igualmente calculados em função dos DRGs e do DS. No Anexo IV podem ser consultados os respectivos valores.

Figura 11
Índices de Casemix segundo os DRGs e o Disease Staging por Hospital
Casos Médicos



Para os casos médicos é o Disease Staging que atribui para a grande maioria dos hospitais (80) Índices de Casemix mais elevados. No entanto, deve referir-se que existe uma associação forte entre os dois sistemas de classificação de doentes, visto que o coeficiente de correlação é de 0.601 (significativo a 0.01).

À semelhança do que foi observado para o total de doentes tratados e para os casos cirúrgicos, nos casos médicos existe uma maior homogeneidade no

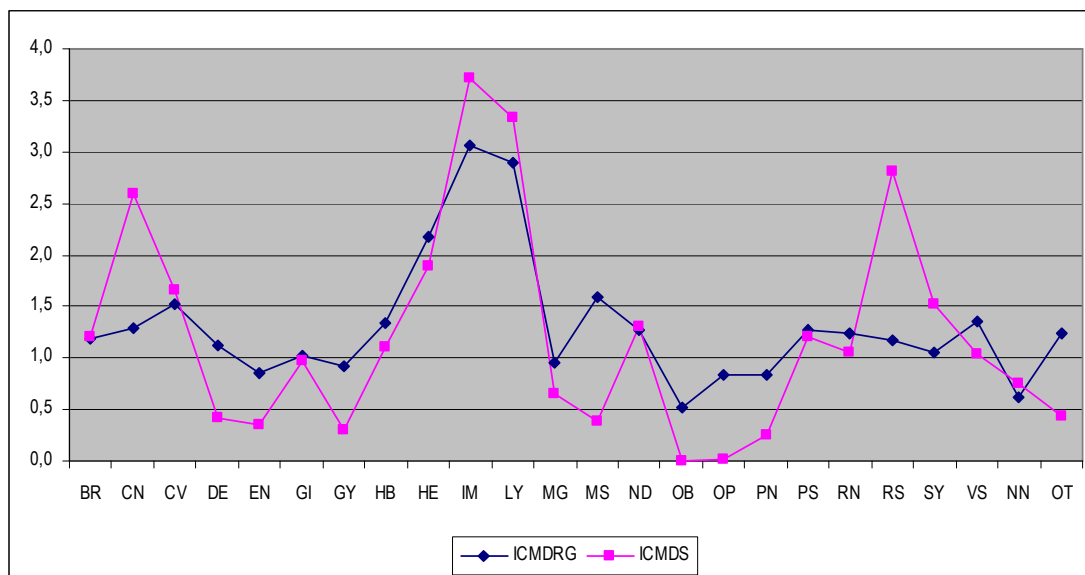
comportamento dos hospitais quando estes são avaliados pelos DRGs, sendo o coeficiente de variação de 0.26, enquanto que com o Disease Staging esta estatística é de 0.42.

Os hospitais com maior Índice de Casemix relativos aos casos médicos e medidos pelos DRGs são os números 11, 88, 87, 86 e 4, enquanto que com o Disease Staging tal é encontrado nos hospitais 81, 73, 11, 65 e 79. Nos 10 hospitais com Índices de Casemix mais elevados pelos dois sistemas de classificação de doentes, somente 3 são comuns.

Para os 8 hospitais em que os DRGs atribuíram Índices de Casemix mais elevados do que o Disease Staging, 5 são do Tipo II e 1 dos Tipos I, III e V.

Para a análise por doenças (observem-se os valores da Figura 12) vai ser utilizado o Grande Agrupamento de Doenças (GADs) do Disease Staging (ver Anexo I para visualizar as siglas correspondentes às doenças). No Anexo IV podem ser consultados os valores.

Figura 12
 Índices de Casemix segundo os DRGs e o Disease Staging por Doença



Para os 24 Grandes Agrupamentos de Doenças (GADs) considerados o Disease Staging atribui Índices de Casemix mais elevados para nove, enquanto que nos restantes tal ocorre com os Diagnosis Related Groups.

Nos nove GADs em que o Disease Staging atribui valores mais elevados comparativamente aos DRGs, as maiores discrepâncias são encontradas nas Doenças respiratórias, nas Doenças do sistema nervoso central e nas Doenças multisistémicas.

Na situação contrária, ou seja Índices de Casemix mais elevados nos DRGs do que no Disease Staging, os valores mais afastados correspondem aos Casos obstétricos, às Doenças oftalmológicas e às Doenças do sistema musculoesquelético.

As situações onde existe uma maior aproximação no Índice de Casemix entre os dois sistemas de classificação de doentes são nas Doenças da mama e nas Doenças endócrinas e metabólicas.

Para caracterizar a gravidade dos doentes vão ser utilizados os três estádios ordinais do Disease Staging. No Quadro XXI são apresentados os Índices de Casemix por nível de severidade e em função dos dois sistemas de classificação de doentes.

Quadro XXI
Índices de Casemix segundo os DRGs e o Disease Staging por Gravidade

Gravidade	DRGs	Disease Staging
Menos Grave	0.96	0.52
Intermédia	1.38	0.92
Mais Grave	1.64	4.04

Por nível de gravidade, enquanto que os DRGs apontam para uma razão entre os doentes mais e menos graves de 1.71 no Índice de Casemix, o Disease Staging apresenta uma razão de 7.77 para os mesmos grupos de doentes.

Também é interessante de referir que para os doentes menos graves e de Gravidade Intermédia os Diagnosis Related Groups apuram Índices de Casemix mais elevados que o Disease Staging, ocorrendo exactamente o contrário nos doentes mais graves.

Para o cálculo do Índice de Casemix por idade, são considerados quatro escalões (até aos 14 anos; 15-44 anos; 45-64 anos e 65 ou mais anos), estando a situação retratada no Quadro XXII.

Quadro XXII
Índices de Casemix segundo os DRGs e o Disease Staging por Idade

Escalão Etário	DRGs	Disease Staging
0-14	0.77	0.12
15-44	0.89	0.20
45-64	1.32	0.84
≥ 65	1.34	2.22

Em qualquer dos sistemas de classificação de doentes o Índice de Casemix aumenta com a idade. Contudo, enquanto que nos DRGs a razão entre valores extremos é de 1.74, no Disease Staging este valor é de 18.5.

Para os escalões etários mais novos (até aos 64 anos) os Diagnosis Related Groups apuram Índices de Casemix mais elevados que o Disease Staging, acontecendo a situação inversa somente nos doentes com 65 ou mais anos.

Assim, em termos de síntese pode afirmar-se que os dois sistemas de classificação de doentes estão a dar informação completamente diferente sobre a produção dos hospitais, quando esta é medida pelo Índice de Casemix.

Esta conclusão é válida para qualquer das perspectivas de análise consideradas, com as seguintes particularidades:

- O Disease Staging identifica maiores diferenças no “casemix” dos hospitais, quando estes são identificados pelo respectivo Tipo;
- Diferenças no “casemix” regional em função do respectivo sistema de classificação de doentes, embora com uma atenuação na razão entre valores extremos, quando comparada com a existente por Tipo de hospital;
- Identificação e ordenação completamente distinta por hospital em função do respectivo sistema de classificação de doentes, inclusivamente com hospitais que passam do “top ten” do Índice de Casemix calculado pelos DRGs, para hospitais com o Índice de Casemix mais baixo, quando calculado pelo Disease Staging. Bons exemplos ilustrativos desta situação são os hospitais 17 e 22;
- Valorização distinta entre os DRGs e o Disease Staging para os casos cirúrgicos e médicos, com o primeiro sistema de classificação de doentes a apurar Índices de Casemix mais elevados para os casos cirúrgicos e o Disease Staging para os casos médicos;
- Valorização completamente distinta do Índice de Casemix por Agrupamento de Doenças, com a particularidade de as Doenças que ocupam as 3^a, 4^a e 7^a posições no Disease Staging, respectivamente as Doenças respiratórias, Doenças do sistema nervoso central e Doenças multisistémicas, apresentarem valores completamente distintos quando medidas pelos DRGs. Ou ainda as Doenças do sistema musculoesquelético, as quais ocupam a 4^a posição com os DRGs e a 19^a com o Disease Staging;
- Menores diferenças, tanto por gravidade dos doentes, como por idade nos DRGs do que no Disease Staging. Ou seja, os Diagnosis Related Groups admitem uma maior homogeneidade entre doentes relativamente a estas características do que o Disease Staging.

Paralelamente, embora se tenha presente o defendido por Hornbrook (1985), o qual afirma que um sistema de classificação de doentes apresenta melhor desempenho para a finalidade para que foi construído, é natural esperar que os dois sistemas possam ser utilizados na avaliação da eficiência dos hospitais.

Assim, para além do valor intrínseco do Índice de Casemix, a sintetização num único indicador da actividade hospitalar, interessa igualmente discutir o seu âmbito de aplicação.

Quando se considera que o financiamento é uma prioridade, então resulta lógico considerar que a informação disponibilizada pelos Diagnosis Related Groups (DRGs) é mais válida, embora se saiba que tem menor validade para avaliar a qualidade dos cuidados prestados.

Pelo contrário, quando se privilegia a qualidade dos cuidados prestados, é igualmente natural deduzir que o Disease Staging apresenta um melhor desempenho, embora se perca algum poder explicativo sobre os custos e concomitantemente sobre os preços de pagamento.

Para tentar resolver esta dicotomia, existem basicamente duas alternativas:

- Utilizar o melhor sistema de classificação de doentes e concomitantemente um Índice de Casemix para cada finalidade;
- Ou, utilizar um Índice de Casemix Composto para responder a todas as finalidades.

Numa primeira abordagem a alternativa que considera a utilização do melhor sistema de classificação de doentes em função da finalidade que se pretende atingir, parece mais atractiva, essencialmente porque desta forma se optimiza a avaliação do desempenho dos hospitais.

Contudo, tanto por motivos estruturais, nomeadamente a existência de duas linhas de autoridade nos hospitais, como por motivos operacionais, dentro dos quais se devem destacar os factos de os sistemas disponibilizarem informações distintas sobre a actividade hospitalar e de valorizarem de forma diferente os produtos destas organizações de saúde (no limite pode afirmar-se que os DRGs estão mais preocupados com aspectos relacionados com a oferta, enquanto que o Disease Staging valoriza mais aspectos da procura), deve questionar-se aquele princípio.

Em primeiro lugar, porque pode contribuir para fracturar ainda mais as duas linhas de autoridade existentes nos hospitais, visto que assume que cada linha pode e deve ter um sistema de informação próprio que, inclusivamente podem conter metodologias de avaliação distintas e eventualmente contraditórias.

Em segundo lugar, porque não existe nenhuma evidência científica que o melhor desempenho de cada sistema de classificação de doentes para a respectiva finalidade (medido por técnicas estatísticas) não contenha erros grosseiros, ou que, no mínimo, disponibilize informações enviesadas sobre o funcionamento dos hospitais.

Neste sentido, opta-se pela segunda alternativa – criação de um Índice de Casemix Composto – o qual apresenta como principais vantagens o de sintetizar informação respeitante às duas linhas de autoridade e de atenuar os valores extremos existentes em cada sistema de classificação de doentes.

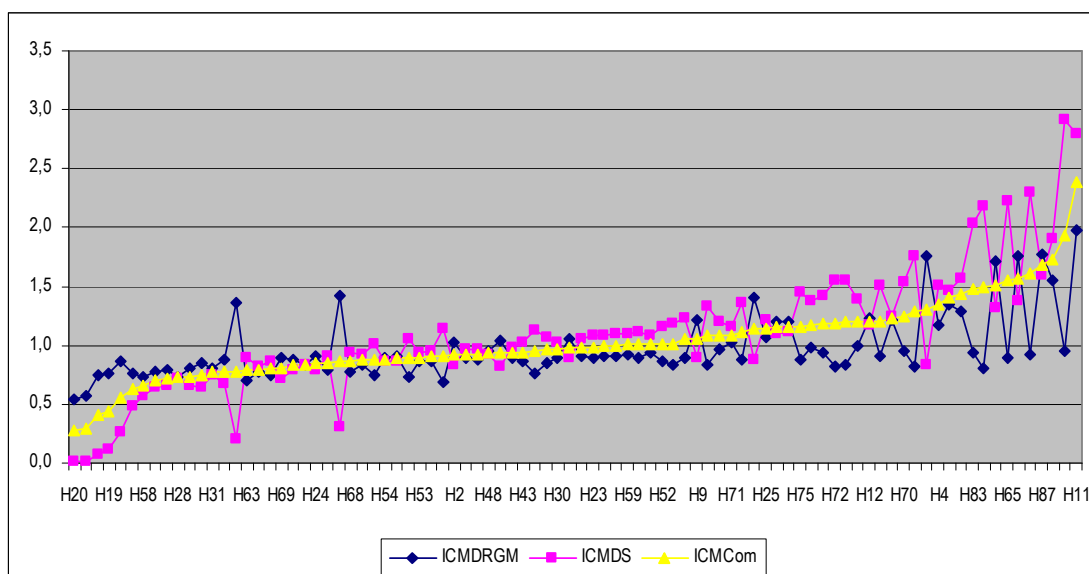
O principal inconveniente poderá ser encontrado numa eventual perda de precisão, dado que a metodologia de construção é mais genérica.

Neste estudo, atendendo aos sistemas de classificação de doentes que estão a ser utilizados, vai ser adoptada a seguinte metodologia para identificar o Índice de Casemix Composto:

- Índice de Casemix dos DRGs (ICMDRG) – Divide-se o valor original de cada hospital pelo Índice de Casemix Nacional dos DRGs;
- Índice de Casemix do Disease Staging (ICMDS) – Divide-se o valor original de cada hospital pelo Índice de Casemix Nacional do DS;
- Índice de Casemix Composto (ICMCom) = (ICMDRG + ICMDS) / 2.

Na Figura 13 é apresentado o Índice de Casemix Composto por hospital, estando estes ordenados crescentemente pelo Índice de Casemix Composto. No Anexo IV podem ser consultados os respectivos valores.

Figura 13
Índice de Casemix Composto por Hospital



O novo Índice de Casemix (ICMCom) encontra-se fortemente correlacionado, tanto com o Índice apurado pelos DRGs (0.673), como com o do Disease Staging (0.923), ambos significativos a 0.01. A correlação entre os Índices de Casemix dos dois sistemas de classificação de doentes é de 0.34.

Os três hospitais com o Índice de Casemix Composto mais elevado são os números 11, 73 e 6, enquanto que os que apresentam o Índice mais baixo são os números 20, 18 e 21.

Nos dez hospitais com Índice de Casemix Composto mais elevado, cinco são do Tipo IV, três do Tipo I e dois do Tipo V. Destes, 5 são comuns com a ordenação proporcionada pelos DRGs e 8 com a do Disease Staging. Deve referir-se que o hospital 5, o qual ocupava a 3^a posição no Índice de Casemix calculado pelos DRGs e a 66^a posição quando calculado pelo Disease Staging, ocupa agora a 14^a posição, enquanto que o hospital 67 ocupava a 8^a posição no Índice de Casemix calculado pelo Disease Staging e a 67^a posição quando se utilizavam os DRGs, ocupa agora a 15^a posição.

Nos 10 hospitais com Índice de Casemix mais baixo, quatro são do Tipo II e três do Tipo III e Tipo IV. Destes 5 são comuns com a ordenação dos DRGs e 7 com a ordenação do Disease Staging. Os hospitais em que se observou uma maior alteração nas respectivas ordenações são os números 17, 22, 45 e 76.

Assim, como comentário final deve dizer-se que face, ao actual estado da arte da medição da produção hospitalar, esta metodologia para se calcular o Índice de Casemix parece mais adequada, visto que incorpora simultaneamente e com o mesmo peso relativo informação sobre a complexidade e sobre a gravidade dos casos tratados em cada hospital.

Contudo, atendendo à importância do assunto sugere-se o seguinte:

- A disponibilização aos hospitais dos três indicadores;
- A escolha dos melhores sistemas de classificação de doentes para serem aplicados aos hospitais portugueses;
- A investigação sobre a melhor metodologia para se atribuírem ponderações para cálculo do Índice de Casemix Composto.

5.1.2. Perfil das Admissões

Conforme foi referido o perfil das admissões dos hospitais deve constituir uma preocupação dominante na administração e na regulação em saúde, não só pelas consequências que pode ter sobre o desempenho destas organizações de saúde, mas também por questões mais relacionadas com a política de saúde, onde avultam as dimensões acessibilidade e disponibilidade dos hospitais.

Neste sentido as interrogações que se colocam são as seguintes:

- Existe ou não evidência de políticas diferentes na admissão de doentes por parte dos hospitais?
- Na eventualidade de tal acontecer quais os aspectos que melhor as justificam? A localização dos hospitais? O tipo de hospitais? Outros?
- Quais as implicações desta política distinta de admissões?

Para se caracterizar a política de admissões dos hospitais serão consideradas as seguintes dimensões:

- A concentração da produção;
- A complexidade e a gravidade dos casos.

Concentração da Produção

A concentração da produção vai ser medida com recurso a dois indicadores:

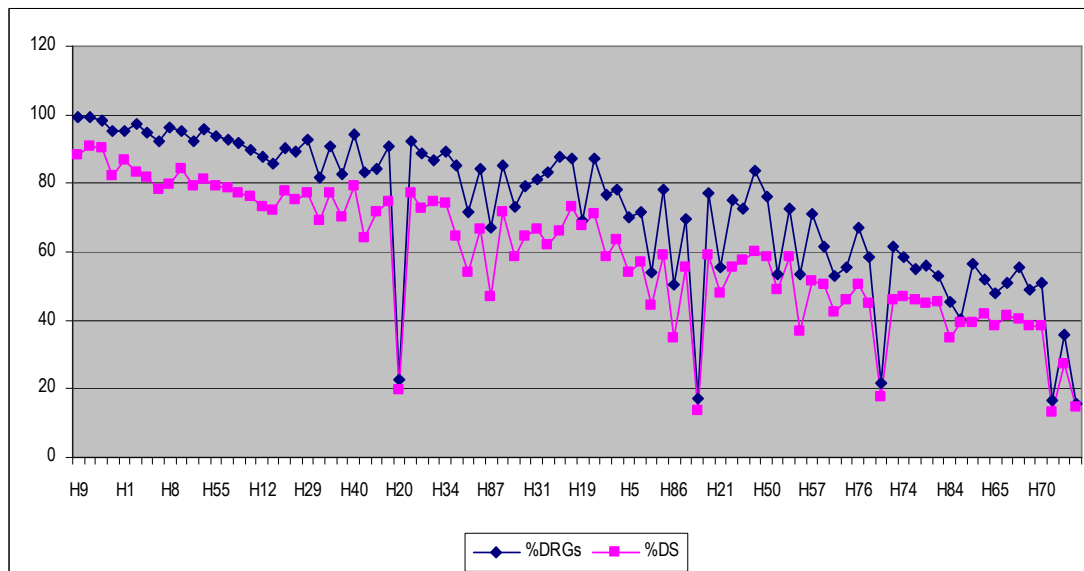
- O número de produtos diferentes em cada hospital;
- O número de produtos responsável por 80% da produção em cada hospital.

Atendendo a que estão a ser utilizados dois sistemas de classificação de doentes, os quais por sua vez disponibilizam informações distintas sobre a actividade dos hospitais (ver capítulo anterior), os indicadores referidos serão calculados para os DRGs e para o Disease Staging.

Aos 2315189 episódios de internamento nos anos de 1999, 2000 e 2001 correspondem 489 produtos diferentes segundo os DRGs e 573 produtos distintos quando se utiliza o Disease Staging.

Na Figura 14 são apresentadas, por hospital, as percentagens de produtos diferentes pelos DRGs e pelo Disease Staging (divisão do número de produtos do hospital pelo número total de produtos identificados em todos os hospitais). Os hospitais estão ordenados decrescentemente pelo total de doentes tratados. Ver Anexo V para se consultarem os respectivos valores.

Figura 14
% de Produtos Diferentes por Hospital segundo
os DRGs e o Disease Staging



Para todos os hospitais, pese embora o facto de para o total dos hospitais existir um maior número de produtos diferentes no Disease Staging do que nos DRGs, a dispersão da produção é maior quando medida por este último sistema de classificação de doentes. A menor diferença é encontrada no hospital 11 (0.88%) e a maior diferença no hospital 42 (23.61%).

Segundo os DRGs os cinco hospitais com maior dispersão da produção são os números 9, 10, 13, 12 e 8 (todos do Tipo I). Com o Disease Staging estes hospitais são os números 10, 13, 9, 1 e 25 (4 do Tipo I e 1 do Tipo III).

Para a maior concentração de produção encontram-se os hospitais 11, 17, 18, 22 e 20 (4 do Tipo II e 1 do Tipo I) de acordo com os DRGs e os hospitais 17, 18, 11, 22 e 20 segundo o Disease Staging.

Apesar das diferenças encontradas entre os dois sistemas de classificação de doentes, tanto na identificação como na ordenação dos hospitais com os maiores e menores níveis de concentração, pode afirmar-se que na globalidade ambos os sistemas estão a disponibilizar o mesmo tipo de informação.

Isto porque, o coeficiente de correlação de Pearson para o nível de concentração da produção entre os DRGs e o Disease Staging é de 0.979 e o coeficiente de correlação de Spearman para as ordenações dos hospitais por nível de concentração é de 0.983.

Contudo, deve ainda referir-se que o *K de Cohen* é de 0.115 o que traduz uma menor concordância entre estes dois sistemas de classificação de doentes na ordenação dos hospitais por nível de concentração da produção.

No entanto, atendendo a que neste particular a questão mais importante em análise é a de apurar o nível de concentração da produção dos hospitais pode concluir-se que é relativamente indiferente a utilização dos DRGs ou do Disease Staging para este efeito, embora este último sistema de classificação apresente sempre níveis de concentração mais elevados.

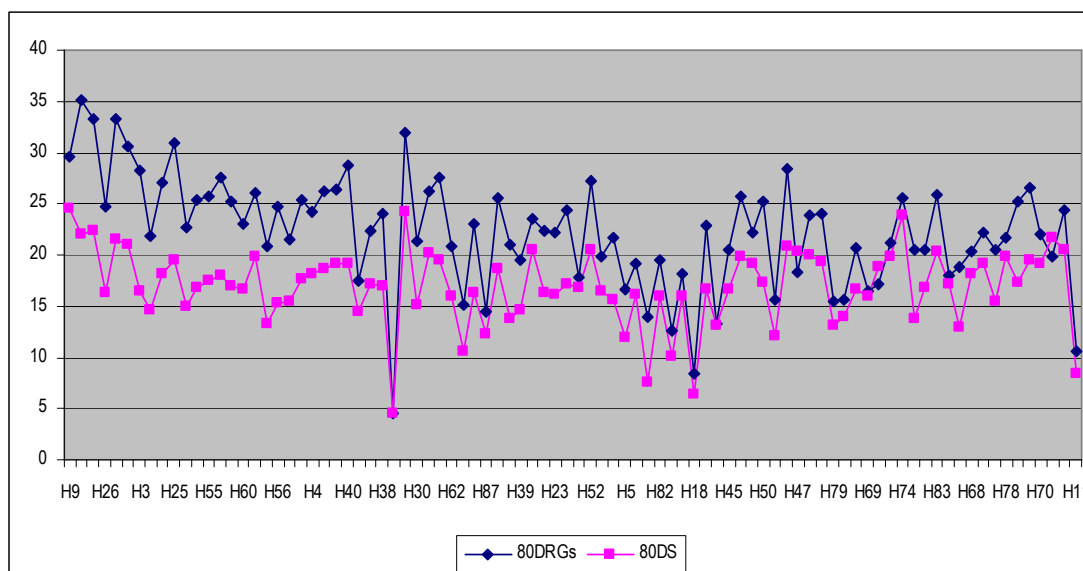
Por outro lado, a concentração da produção dos hospitais medida pelo número de produtos diferentes está fortemente associada com o volume total de produção, visto que o coeficiente de correlação de Pearson entre estas duas variáveis é de 0.678 quando se utilizam os DRGs e de 0.750 com o Disease Staging.

Assim, pode igualmente concluir-se que a dispersão da produção dos hospitais, medida pelo número de produtos diferentes, varia directamente com o aumento da produção. Ou seja, quanto maior a produção mais casos raros existem nos hospitais (efeito raridade).

Neste sentido, é natural que entre os 10 hospitais com maior dispersão da produção, 7 sejam do Tipo I e 3 do Tipo III com os DRGs e com o Disease Staging.

Na Figura 15 são apresentadas, por hospital, as percentagens de produtos diferentes responsáveis por 80% da produção, por DRG e pelo Disease Staging. Os hospitais estão ordenados decrescentemente pelo total de doentes tratados. Ver Anexo V para se consultarem os respectivos valores.

Figura 15
% de Produtos Diferentes Responsáveis por 80% da Produção,
por Hospital, segundo os DRGs e o Disease Staging



Quando a concentração da produção dos hospitais é medida pelo número de produtos diferentes que justificam 80% do total da produção, mais uma vez se encontra uma menor dispersão no Disease Staging do que nos DRGs. Com a particularidade de existirem 3 hospitais (47, 17 e 22) em que a dispersão é maior no Disease Staging. O hospital 10 é aquele em que as diferenças entre o nível de concentração é maior entre os dois sistemas de classificação de doentes (13.08%).

Segundo os DRGs os cinco hospitais com maior dispersão da produção são os números 10, 1, 13, 7 e 25 (4 do Tipo I e 1 do Tipo III), enquanto que com o Disease Staging tal ocorre nos hospitais 9, 7, 74, 13 e 10 (4 do Tipo I e 1 do Tipo IV).

Com maiores níveis de concentração da produção encontram-se os hospitais 20, 18, 11, 86 e 21 (3 do Tipo II e 1 dos Tipos I e V) de acordo com os DRGs e os hospitais 20, 18, 14, 11 e 86 (2 dos Tipos I e II e 1 do Tipo V) segundo o Disease Staging.

Mais uma vez os dois sistemas de classificação de doentes parecem estar a dar a mesma informação sobre o nível de concentração da produção quando se utiliza a percentagem de produtos diferentes responsáveis por 80% da produção. O coeficiente de correlação de Pearson é de 0.823 para o nível de concentração e o coeficiente de correlação de Spearman é de 0.742 para as respectivas ordenações por nível de concentração.

À semelhança do que foi referido anteriormente o *K de Cohen* (0.034) parece traduzir uma concordância fraca entre estes dois sistemas de classificação na ordenação dos hospitais por nível de concentração da produção.

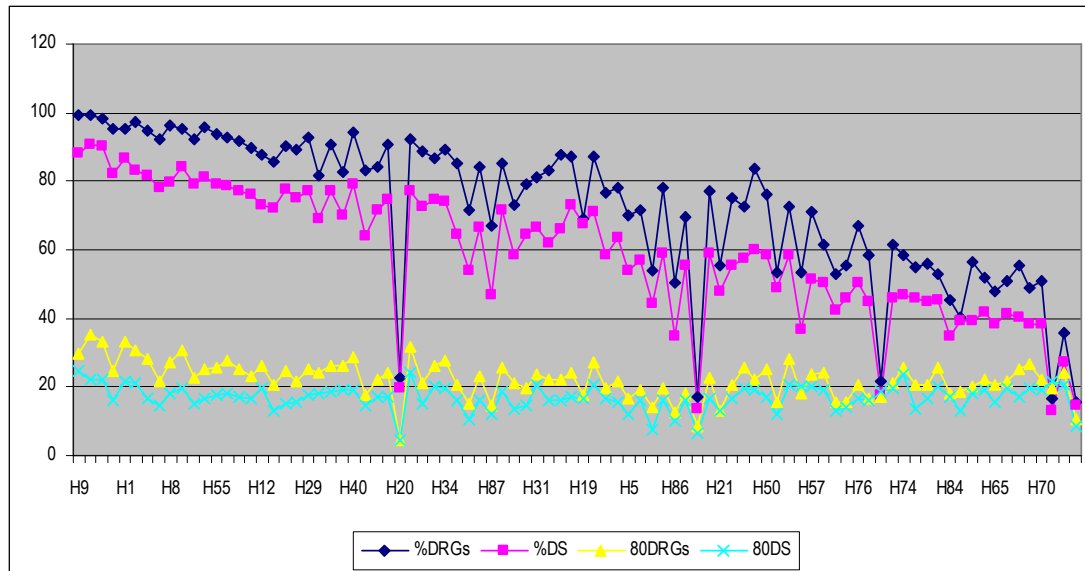
Neste sentido, pode ainda concluir-se que na ausência de outros aspectos os dois sistemas de classificação de doentes parecem estar a disponibilizar a mesma informação sobre a concentração da produção dos hospitais.

Por outro lado, a concentração da produção quando se utiliza este indicador parece estar associada com o volume da produção quando se utilizam os DRGs (o coeficiente de correlação é de 0.526). Quando se utiliza o Disease Staging esta associação é mais baixa, visto que esta estatística toma o valor 0.240, ainda assim significativa a um nível de 0.05.

Daqui pode concluir-se que o efeito volume da produção tem implicações mais atenuadas na concentração da produção, quando esta é medida pelo número de produtos responsáveis por 80% da produção, em oposição à simples enumeração do número total de produtos diferentes em cada hospital. Deve ainda referir-se que este fenómeno é ainda mais saliente quando se utiliza o Disease Staging (ver Figura 16).

Ainda assim dos 10 hospitais com maior dispersão da produção a 80%, 7 são do Tipo I e 3 do Tipo III com os DRGs, embora com o Disease Staging 6 sejam do Tipo I, 2 do Tipo IV e 1 dos Tipos II e III.

Figura 16
 Concentração da Produção por Hospital
 Total de Produtos Diferentes e Produtos Responsáveis por 80% da Produção
 DRGs e Disease Staging



A utilização conjunta destes dois indicadores para a concentração da produção por hospitais permite evidenciar o seguinte:

- Existe uma maior variação no comportamento dos hospitais em função da dimensão da sua produção para o total dos produtos diferentes (efeito raridade) do que para a concentração de produtos a 80% da produção;
- De facto, enquanto que o coeficiente de variação é de 0.29 ou de 0.32, respectivamente para os DRGs e para o Disease Staging, para o total de produtos diferentes, esta estatística, para os produtos responsáveis por 80% da produção, é de 0.24 e de 0.22 para os mesmos sistemas de classificação de doentes;
- Assim, pode concluir-se que a simples tipologia e dimensão dos hospitais, embora exprima com grande intensidade o efeito raridade da produção, não é por si esclarecedora sobre a concentração da produção;
- Finalmente, deve referir-se que embora os dois sistemas de classificação de doentes pareçam estar a disponibilizar o mesmo tipo de informação sobre o comportamento dos hospitais em relação ao seu modo de produção, o Disease Staging parece apresentar uma maior sensibilidade aos efeitos da dimensão da produção, visto que torna menos directa esta relação.

Complexidade e Gravidade dos Casos

A escolha de doentes mais favoráveis por parte dos hospitais pode assumir duas perspectivas:

- Selecção de produtos/doenças mais favoráveis;
- Selecção de situações de saúde mais favoráveis.

Na primeira situação, selecção de produtos por parte dos hospitais, as principais justificações podem ser encontradas na possibilidade de os hospitais competirem por doenças que lhes são mais favoráveis em termos de resultados e concomitantemente no respectivo desempenho.

Até ao presente estudo e no actual estado da arte e atendendo igualmente à realidade existente em Portugal, o melhor mecanismo para se investigarem comportamentos diferentes por parte dos hospitais no respectivo “mix” das doenças tratadas pode ser analisado mediante o recurso à informação disponibilizada pelos DRGs.

Isto porque, atendendo às características dos DRGs, essencialmente no que respeita à definição de produtos com consumo de recursos homogéneo e à existência de um peso relativo para cada DRG, pode afirmar-se que este sistema de classificação de doentes pode traduzir a complexidade dos casos tratados em cada hospital (Hornbrook, 1982; Aronow, 1988; Rosko, 1988; Vladeck e Kramer, 1988 e Costa, 1994).

Para operacionalizar este conceito são utilizados os pesos relativos por DRG em vigor em 2001 (Portaria nº 189/2001 de 9 de Março) para se definirem níveis de complexidade dos produtos dos hospitais.

Para tal foram definidos três níveis de complexidade, de acordo com a seguinte metodologia:

- Identificação dos níveis de complexidade através dos percentis 33.3 e 66.7, para os pesos relativos;
- Atribuição do nível de complexidade 1 a todos os DRGs cujos pesos relativos estejam incluídos até ao percentil 33.3. Com o nível 2 de complexidade são classificados os DRGs com pesos relativos entre os percentis 33.3 e 66.7. Finalmente para o nível 3 são considerados os DRGs cujos pesos relativos sejam superiores ao percentil 66.7.

Assim, os níveis de complexidade são os seguintes:

- Nível 1 (PR1) – DRGs com pesos relativos até 0.82 (162 DRGs);
- Nível 2 (PR2) – DRGs com pesos relativos entre 0.84 e 1.53 (164 DRGs);
- Nível 3 (PR3) – DRGs com pesos relativos superiores a 1.54 (163 DRGs).

Para a segunda situação, selecção de situações de saúde mais favoráveis por parte dos hospitais, vai ser analisada a severidade do estado do doente, igualmente tendo em atenção que o tratamento de doentes menos graves poderá otimizar os resultados dos hospitais, com as naturais consequências no respectivo desempenho.

Para tal serão utilizados os estadios de severidade do Disease Staging, correspondendo genericamente:

- Estadio 1 (ST1) – Doentes menos graves;
- Estadio 2 (ST2) – Doentes com gravidade intermédia;
- Estadio 3 (ST3) – Doentes mais graves.

No Quadro XXIII são apresentados para todos os doentes e casos cirúrgicos e médicos, os doentes tratados por nível de complexidade, bem como por grau de severidade.

Quadro XXIII
Níveis de Complexidade e Graus de Severidade
Total de Doentes, Casos Cirúrgicos e Casos Médicos

	Nível 1	Nível 2	Nível 3	Estadio 1	Estadio 2	Estadio 3
Casos Cirúrgicos	375800	285202	236557	708745	155926	31594
Casos Médicos	747583	553956	116091	877203	291211	234123
Total de Doentes	1123383	839158	352648	1585948	447137	265717

Como primeira observação deve referir-se que para os estadios de gravidade não estão compatibilizados 16387 doentes, dos quais 15093 correspondem a casos médicos e 1294 a casos cirúrgicos. Tal facto deve-se, como foi referido anteriormente, à existência de situações em que o Disease Staging atribui um estadio 0.

Para os níveis de complexidade e em relação ao total de admissões existem mais casos médicos nos níveis 1 e 2, respectivamente de 66.6% e 66%, enquanto que no nível 3 tal ocorre nos casos cirúrgicos (67.1%).

A mesma comparação para os estadios de gravidade permite concluir que são sempre os casos médicos que apresentam maiores percentagens de doentes tratados em cada estadio, respectivamente de 55.3%, 65.1% e 88.1% para os estadios 1, 2 e 3.

Para os casos cirúrgicos, observa-se uma distribuição mais uniforme por níveis de complexidade, do que por estadios de gravidade, visto que para a complexidade o máximo é atingido no nível 1 (41.9%) e o mínimo no nível 3 (26.4%), enquanto que para a severidade, embora o máximo e o mínimo sejam igualmente atingidos nos estadios 1 e 3, estes valores são de, respectivamente 79.1% e 3.5%.

Nos casos médicos o panorama é distinto, embora, tanto para a complexidade, como para a severidade o máximo e o mínimo sejam igualmente atingidos nos graus 1 e 3. Contudo, para a complexidade a razão entre valores extremos é maior, visto que estes valores são de 52,7% e 8.2%, em oposição a 62.5% e 16.7% encontrados na distribuição por estadios de gravidade.

Conforme se observa, estes dois indicadores parecem estar a disponibilizar informações diferentes sobre a actividade dos hospitais, visto que os casos mais complexos são encontrados nos casos cirúrgicos e os casos mais graves nos casos médicos.

Para melhor se visualizar este comportamento, observem-se os valores do Quadro XXIV, nos quais estão indicadas as percentagens de doentes tratados por níveis de complexidade e por estadios de gravidade, igualmente para todos os doentes, casos cirúrgicos e médicos.

Quadro XXIV
% de Doentes Tratados por Níveis de Complexidade e Estadios de Gravidade
Todos os Doentes, Casos Cirúrgicos e Casos Médicos

	Estadio 1	Estadio 2	Estadio 3
Casos Cirúrgicos			
Nível 1	92.45	7.48	0.06
Nível 2	83.42	16.14	0.45
Nível 3	52.57	34.69	12.75
Casos Médicos			
Nível 1	78.38	18.80	2.82
Nível 2	41.74	25.46	32.79
Nível 3	61.44	10.72	27.84
Total de Doentes			
Nível 1	83.14	14.97	1.89
Nível 2	55.91	22.29	21.80
Nível 3	55.48	26.83	17.70

No Quadro XXIV estão identificados a “azul” as células que deveriam ter o valor máximo se existisse concordância entre as duas formas de caracterizar a produção dos hospitais, complexidade e severidade, a “vermelho” os valores máximos realmente observados e a “verde” os que cumprem simultaneamente as duas condições.

Daqui, pode concluir-se que para todas as admissões e para os casos cirúrgicos e médicos, somente existe concordância entre o nível 1 de complexidade e o estadio 1 de gravidade, visto que para os restantes níveis de complexidade (2 e 3) os máximos são sempre igualmente encontrados no estadio 1.

É conveniente ainda destacar a situação dos casos cirúrgicos e para a total de admissões, nos quais se observa que para o nível de complexidade 3, a menor frequência de doentes é encontrada no estadio 3, o que sobreleva ainda mais as diferenças entre complexidade e severidade para se apreciar a actividade dos hospitais.

Estando genericamente identificadas as diferenças entre complexidade e severidade para se caracterizar a actividade dos hospitais, torna-se neste momento importante definir um cenário para se avaliar o perfil das admissões dos hospitais.

Para a complexidade a situação não é unívoca, visto que por um lado pode argumentar-se que uma predominância de doentes menos complexos pode minimizar custos e eventualmente otimizar a qualidade dos cuidados prestados, por outro lado pode igualmente defender-se que a predominância de casos mais complexos pode contribuir para o aumento das receitas dos hospitais e desta forma contribuir para a melhoria do desempenho financeiro dos hospitais e eventualmente da eficiência.

No que se refere à severidade o panorama é mais linear, visto que sem mecanismos de ajustamento pelo risco, os hospitais que à partida admitam doentes menos graves, apresentam melhores condições para otimizar a eficiência e a qualidade dos cuidados prestados, atendendo a que este tipo de doentes potencialmente conduz a resultados mais favoráveis.

Inclusivamente no plano do financiamento e do desempenho financeiro dos hospitais, esta relação é igualmente directa, essencialmente atendendo ao facto de o sistema de financiamento actualmente em vigor não contemplar a gravidade dos doentes, a predominância de situações menos graves em determinado hospital pode conduzir a um menor risco financeiro.

Face ao exposto, a política dos hospitais no que respeita à escolha de doentes vai ser caracterizada da seguinte forma:

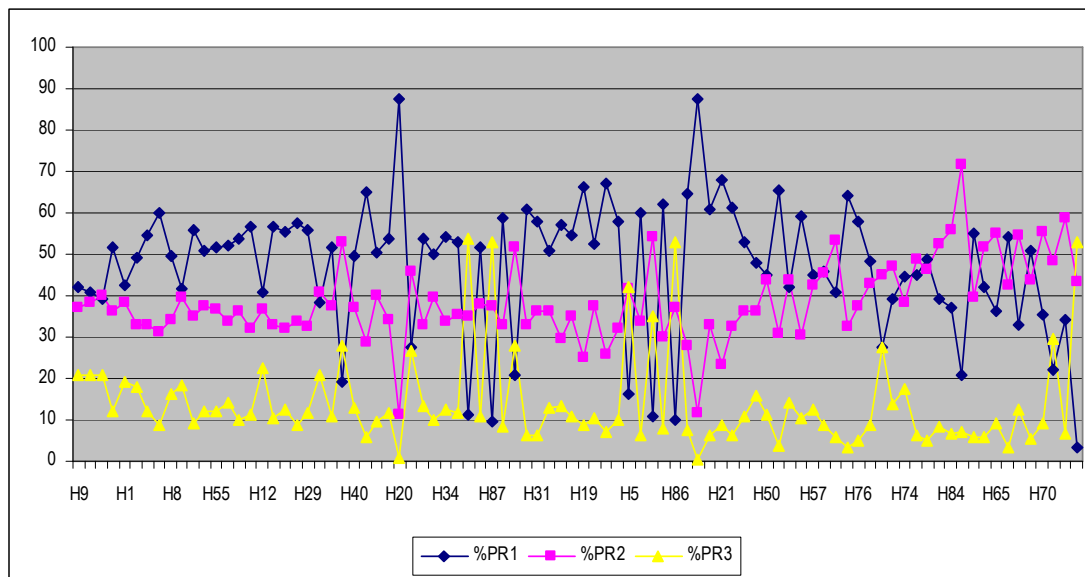
- Identificação de eventuais comportamento distintos por parte dos hospitais no que se refere à complexidade e à gravidade dos casos tratados;
- Comparação do desempenho dos hospitais em função da complexidade e da severidade, assumindo-se que as organizações de saúde que apresentem tendências mais discrepantes são aquelas que potencialmente praticam uma maior escolha de doentes.

Complexidade dos casos

A caracterização da complexidade dos doentes tratados por hospital vai ser feita de acordo com a metodologia referida e será apresentada para todos os doentes tratados e para os casos cirúrgicos e médicos.

Na Figura 17 são apresentadas para todos os doentes tratados, a percentagem de doentes admitidos por nível de complexidade, estando os hospitais ordenados decrescentemente pela sua produção. Ver Anexo VI para se consultarem os respectivos valores.

Figura 17
% de Doentes por Níveis de Complexidade
Total de Admissões



Em relação ao nível de complexidade mais baixo observam-se os seguintes aspectos:

- Os hospitais 20, 18 e 21 são os que apresentam mais doentes neste nível. Nos 10 hospitais com maiores percentagens encontram-se 4 do tipo II e 3 dos Tipos III e IV;
- Os hospitais com menos admissões neste nível são os números 11, 87 e 86. Para os 10 hospitais com menores percentagens de admissões 5 são do Tipo I, 3 do Tipo V e 1 dos Tipos II e IV.

Para o nível de complexidade 2 a situação é a seguinte:

- Os hospitais 73, 81 e 84 são os que tratam mais doentes. Nos 10 hospitais com maiores percentagens de admissões 8 são do Tipo IV e 2 do Tipo I;
- Com menos admissões encontram-se os hospitais 20, 18 e 21. Nos 10 hospitais com menores percentagens de admissões 4 são dos Tipos II e III e 2 do Tipo IV.

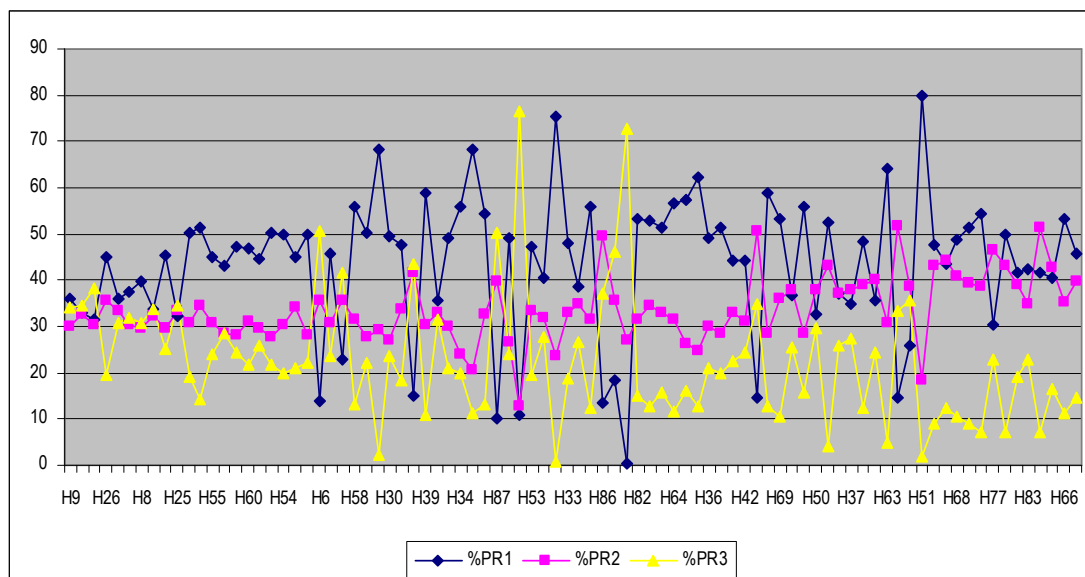
Para o nível de complexidade mais elevado a situação é a seguinte:

- Os hospitais 88, 11 e 87 são os que apresentam maiores admissões. Nos 10 hospitais com maiores percentagens 5 são do Tipo I, 3 do Tipo V e 2 do Tipo II;
- Com menos admissões encontram-se os hospitais 18, 20 e 78. Nos 10 hospitais com menores percentagens de admissões 7 são do Tipo IV, 2 do Tipo II e 1 do Tipo III.

O comportamento dos hospitais no que respeita às admissões em função da respectiva complexidade é heterogéneo entre os 3 níveis considerados. De facto, o coeficiente de variação é de 0.33, 0.24 e 0.80, respectivamente para os níveis de complexidade 1, 2 e 3, pelo que se pode afirmar que a variação entre hospitais é mais expressiva nos casos mais complexos.

Na Figura 18 são apresentadas para os casos cirúrgicos, a percentagem de doentes admitidos por nível de complexidade, estando os hospitais ordenados decrescentemente pela sua produção. Ver Anexo VI para se consultarem os respectivos valores.

Figura 18
% de Doentes por Níveis de Complexidade
Casos Cirúrgicos



Para os casos com menor complexidade observa-se o seguinte:

- Os hospitais 51, 18 e 20 são os que apresentam mais admissões neste nível. Nos 10 hospitais com maiores percentagens de admissões 5 são do Tipo III, 3 do Tipo II e 2 do Tipo IV;
- Com menos doentes encontram-se os hospitais 14, 87 e 5. Nos 10 hospitais nesta situação 5 são do Tipo I, 3 do Tipo V e 2 do Tipo II.

Para o nível de complexidade 2 podem fazer-se os seguintes comentários:

- Os hospitais 17, 78 e 22 são os que apresentam mais admissões. Considerando os 10 hospitais observa-se que 7 são do Tipo IV, 2 do Tipo II e 1 do Tipo V;
- Com menos admissões encontra-se os hospitais 5, 51 e 21. Para os 10 hospitais com menores percentagens de admissões 7 são do Tipo III, 2 do Tipo II e 1 do Tipo I.

Para os casos mais complexos destacam-se os seguintes aspectos:

- Os hospitais 5, 14 e 6 são os que apresentam mais admissões, observando-se para os 10 hospitais nesta situação que 6 são do Tipo I, 3 do Tipo V e do Tipo IV;
- Os hospitais 18, 51 e 20 são os que apresentam menos admissões neste nível de complexidade. Para os 10 hospitais com menores percentagens de admissões 7 são do Tipo IV, 2 do Tipo II e 1 do Tipo III.

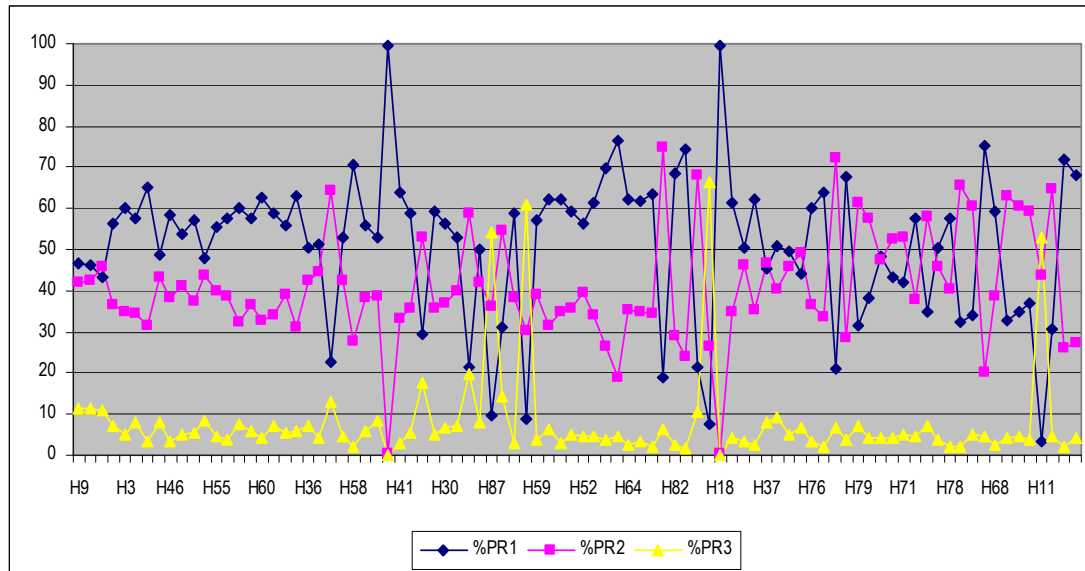
À semelhança do referido para o total das admissões, o comportamento dos hospitais, nos casos cirúrgicos, é heterogéneo entre os 3 níveis de complexidade, assumindo uma maior expressão no nível 3 e uma menor no nível 2, sendo os coeficientes de variação de 0.34, de 0.20 e de 0.59, respectivamente para os níveis de complexidade 1, 2 e 3.

Na Figura 19 são apresentadas para os casos médicos, a percentagem de doentes admitidos por nível de complexidade, estando os hospitais ordenados decrescentemente pela sua produção. Ver Anexo VI para se consultarem os respectivos valores.

Para o nível 1 de complexidade observa-se o seguinte:

- Os hospitais 20, 18 e 19 são os que apresentam mais admissões. Para os 10 hospitais nesta situação 5 são do Tipo II, 3 do Tipo III e 2 do Tipo IV;
- Com menos admissões encontram-se os hospitais 11, 86 e 88. Quando se consideram os 10 hospitais com menores percentagens de admissões 6 são do Tipo I, 3 do Tipo V e 1 do Tipo IV.

Figura 19
 % de Doentes por Níveis de Complexidade
 Casos Médicos



Para o nível 2 de complexidade destacam-se os seguintes aspectos:

- Os hospitais 14, 73 e 5 são os que apresentam mais admissões. Nos 10 hospitais nesta situação 7 são do Tipo IV e 3 do Tipo I;
- Com menos admissões encontram-se os hospitais 20, 18 e 19. Nos 10 hospitais com menos admissões 5 são do Tipo II, 3 do Tipo III e 1 dos Tipos IV e V.

Para os casos mais complexos podem fazer-se os seguintes comentários:

- Os hospitais 86, 88 e 87 são os que apresentam mais admissões. Nos 10 hospitais nesta situação 7 são do Tipo I e 3 do Tipo V;
- Com menos admissões encontram-se os hospitais 18, 20 e 85. Nos 10 hospitais nesta situação 5 são do Tipo IV, 3 do Tipo II e 2 do Tipo III.

Nos casos médicos, o comportamento entre hospitais nas admissões por níveis de complexidade é genericamente semelhante ao referido para o total das admissões e para os casos cirúrgicos. Contudo, deve evidenciar-se que o coeficiente de variação é relativamente baixo e similar para os níveis de complexidade 1 e 2 (0.35 e 0.33), enquanto que para o nível 3 de complexidade assume um valor mais elevado (1.46).

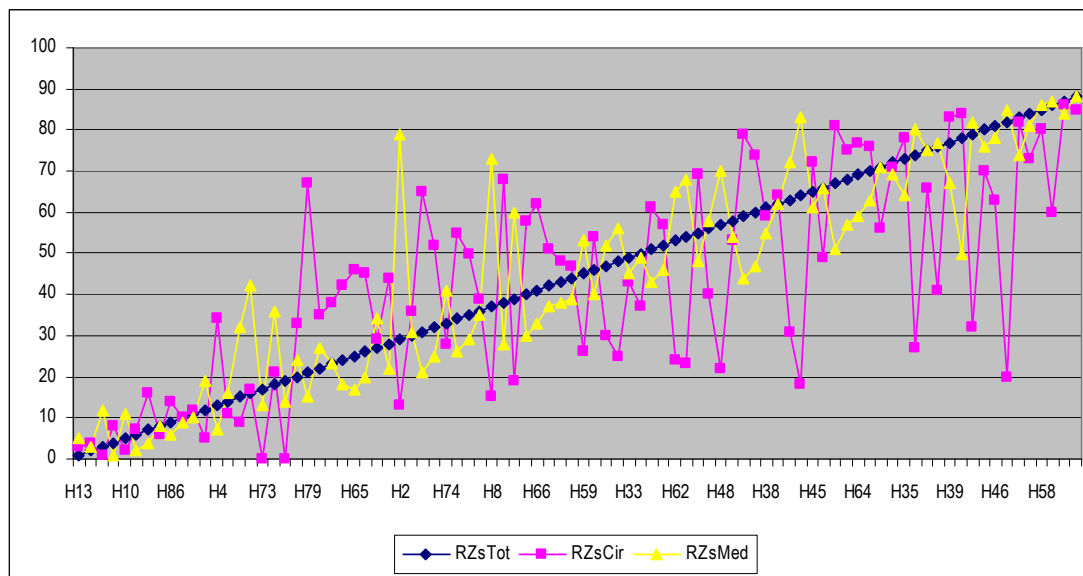
Em seguida, a complexidade das admissões dos hospitais será sintetizado num indicador único. Para tal será utilizado o “z score” de acordo com a equação descrita na Metodologia.

Na Figura 20 são apresentadas as ordenações dos hospitais, para o total de doentes e casos cirúrgicos e médicos.

Para o total de admissões observa-se o seguinte:

- O hospital com um “z score” mais baixo (menos doentes do nível de complexidade 1) foi o número 13, enquanto que o hospital 20 apresenta o valor mais elevado para este indicador (mais doentes do nível de complexidade 1);
- Para os 10 hospitais com um “z score” mais baixo 7 são do Tipo I e 3 do Tipo V, estando 6 localizados na Região de Saúde IV e 2 nas Regiões III e V;
- Nos 10 hospitais com um “z score” mais elevado 6 são do Tipo III, 3 do Tipo II e 1 do Tipo I.

Figura 20
Ordenação dos Hospitais por Complexidade (“Z score”)
Total de Doentes e Casos Cirúrgicos e Médicos



Nos casos cirúrgicos destacam-se os seguintes aspectos:

- O hospital 9 apresenta o “z score” mais baixo, enquanto que o hospital 18 apresenta o valor mais elevado;
- Nos 10 hospitais com “z score” mais baixo 8 são do Tipo I e 2 do Tipo V, estando 6 localizados na Região IV, e na Região V e 1 na Região III;

- Para os 10 hospitais com “z score” mais elevado 6 são do Tipo III, 3 do Tipo II e 1 do Tipo IV, estando 7 localizados na Região de Saúde V, 2 na Região III e 1 na Região IV.

Para os casos médicos a situação é a seguinte:

- O hospital 87 apresenta o “z score” mais baixo, enquanto que o hospital 20 apresenta o valor mais elevado;
- Nos 10 hospitais com “z score” mais baixo 7 são do Tipo I e 3 do Tipo V, estando 8 localizados na Região de Saúde IV e 1 nas Regiões III e V;
- Nos 10 hospitais com “z score” mais elevado 5 são do Tipo III, 3 do Tipo II e 2 do Tipo I, estando 5 localizados na Região de Saúde V, 4 na Região IV e 1 na Região III.

Em conclusão pode afirmar-se que se observam comportamentos distintos entre os casos médicos e cirúrgicos, sendo o *K de Cohen* de 0.00. A não concordância na ordenação dos hospitais mantêm-se entre o total de admissões e os casos cirúrgicos (*K de Cohen* é igual 0.012) e entre o total de admissões e os casos médicos (*K de Cohen* é igual 0.082).

Deve igualmente referir-se que parece não existir qualquer padrão regional na complexidade dos doentes tratados, sendo frequente encontrar hospitais pertencentes à mesma Região de Saúde, com maiores e menores proporções na complexidade dos doentes admitidos.

Para o Tipo de hospital, pesem embora as diferenças apontadas entre os casos cirúrgicos e os médicos, são na generalidade os hospitais do Tipo I, aqueles que tratam doentes mais complexos, enquanto que na situação contrária, são encontrados os hospitais do Tipo III como aqueles que tratam doentes menos complexos.

Gravidade dos casos

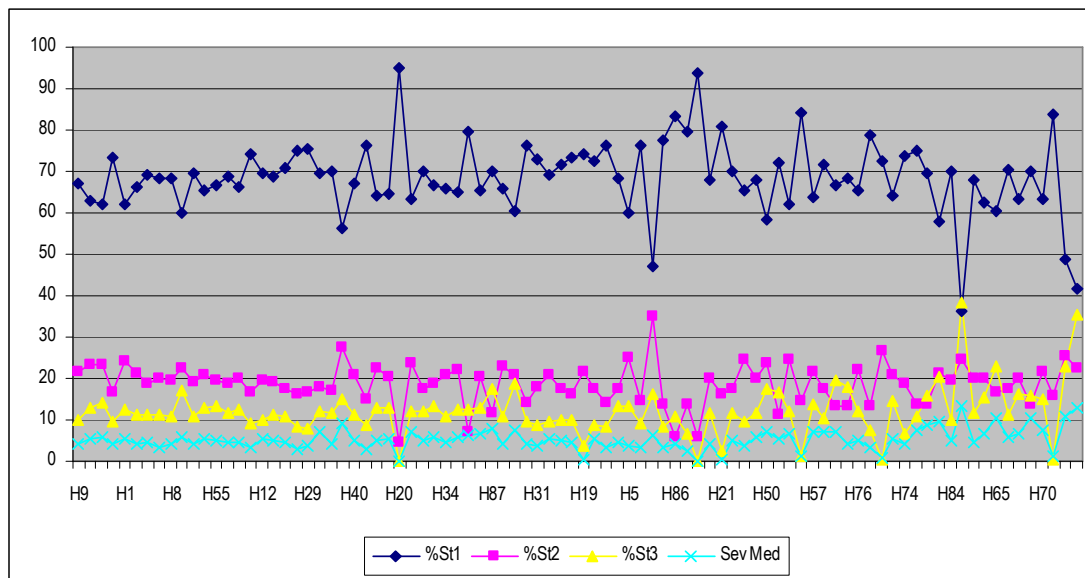
Esta dimensão pode ser medida de duas formas, a distribuição dos doentes pelos estadios ordinais de severidade do Disease Staging, ou pelo valor médio da gravidade, tal como foi calculado para a mortalidade esperada.

Para este tipo de estudo, embora o segundo indicador seja por vezes utilizado, dá-se preferência à distribuição dos doentes pelos escalões de gravidade, como a medida que melhor pode exprimir a política de admissões de cada hospital.

Esta análise será feita para a globalidade dos doentes, bem como para os casos cirúrgicos e para os casos médicos.

Na Figura 21 são apresentadas, por hospital, as percentagens de doentes admitidos nos estadios de severidade 1 (menos grave), 2 (intermédio) e 3 (mais grave), bem como a gravidade média dos doentes, para todas as admissões. Os hospitais estão ordenados decrescentemente pela sua produção. Ver Anexo VII para se consultarem os respectivos valores.

Figura 21
% de Doentes por Estadios de Severidade e Severidade Média por Hospital
Todas as Admissões



Em relação ao nível de gravidade mais baixo (estadio 1) podem fazer-se os seguintes comentários:

- Os hospitais que apresentam mais doentes neste estadio são os números 20, 18 e 47. Entre os 10 hospitais com maiores percentagens de doentes, encontram-se 4 do Tipo II, 3 do Tipo IV, 2 do Tipo V e 1 do Tipo II;
- Os hospitais que apresentam menos doentes neste estadio são os números 73, 11 e 14. Nos 10 hospitais com percentagens mais baixas de doentes, 5 são do Tipo I, 3 do Tipo IV e IV e 2 do Tipo 3.

Para os doentes do estadio 2 destacam-se os seguintes aspectos:

- Os hospitais 14, 6 e 22 são os que apresentam maiores frequências relativas deste tipo de doentes. Para os 10 hospitais com maiores percentagens, 5 são do Tipo I, 2 dos Tipos III e IV e 1 do Tipo II;
- Os hospitais 20, 86 e 18 são os que apresentam menores percentagens de doentes tratados no estadio 2. Considerando os 10 hospitais nesta categoria 4 são do Tipo IV, 3 do Tipo V, 2 do Tipo II e 1 do Tipo III.

No que se refere aos doentes mais graves, a situação é a seguinte:

- Os hospitais 73, 11 e 81 são os que tratam mais doentes mais graves. Entre os 10 hospitais que tratam doentes mais graves encontram-se 6 do Tipo IV, 2 do Tipo I e 1 dos Tipos III e V;
- Com menos doentes neste estadio de severidade estão os hospitais 20, 18 e 22. Nos 10 hospitais nesta situação, 6 são do Tipo II, 3 do IV e 1 do III.

Para a severidade média, podem tecer-se os seguintes comentários:

- Os hospitais 73, 11 e 81 são os que tratam doentes com maior gravidade média. Nos 10 hospitais nesta situação, 7 são do Tipo IV, 2 do I e 1 do V;
- Com menor gravidade média encontram-se os hospitais 18, 20 e 21. Nos 10 hospitais com menor gravidade média em todos os doentes tratados, encontram-se 6 do Tipo II, 3 do III e 1 do IV.

Por outro lado, a variabilidade entre os hospitais aumenta de uma forma directamente proporcional com o nível de gravidade dos doentes, visto que os coeficientes de variação são de 0.13, 0.26 e 0.48, respectivamente para os estadios 1, 2 e 3. Deve ainda ser referido que o coeficiente de variação para a severidade média é de 0.47.

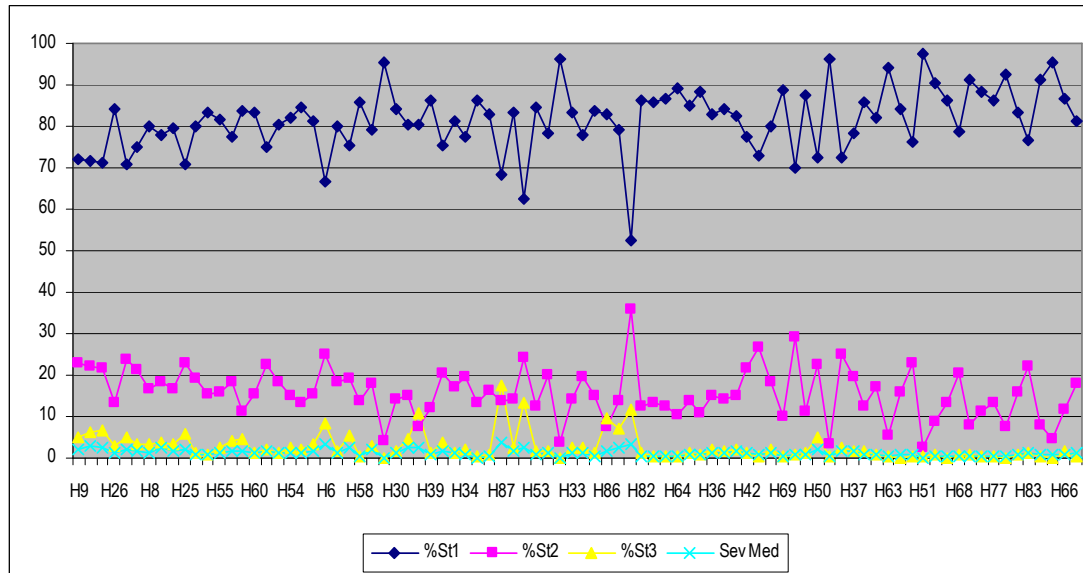
Assim, pode afirmar-se que existem comportamentos distintos dos hospitais em função da gravidade dos doentes, bem mais patente nos doentes mais graves (estadio 3) e na severidade média por hospital.

Na Figura 22 são apresentadas, por hospital, as percentagens de doentes admitidos nos estadios de severidade 1 (menos grave), 2 (intermédio) e 3 (mais grave), bem como a gravidade média dos doentes, para os casos cirúrgicos. Os hospitais estão ordenados decrescentemente pela sua produção. Ver Anexo VII para se consultarem os respectivos valores.

Em relação ao nível de gravidade mais baixo (estadio 1) podem fazer-se os seguintes comentários:

- Os hospitais que apresentam mais doentes neste estadio são os números 51, 79 e 18. Entre os 10 hospitais com maiores percentagens de doentes, encontram-se 7 do Tipo IV, 2 do Tipo II e 1 do Tipo III;
- Os hospitais que apresentam menos doentes neste estadio são os números 14, 5 e 6. Nos 10 hospitais com percentagens mais baixas de doentes, 7 são do Tipo I, 2 do Tipo III e 1 do Tipo V.

Figura 22
% de Doentes por Estádios de Severidade e Severidade Média por Hospital
Casos Cirúrgicos



Para os doentes do estadio 2 destacam-se os seguintes aspectos:

- Os hospitais 14, 24 e 22 são os que apresentam maiores frequências relativas deste tipo de doentes. Para os 10 hospitais com maiores percentagens, 5 são do Tipo I, 3 do Tipo III e 1 dos Tipos II e IV;
- Os hospitais 51, 79 e 18 são os que apresentam menores percentagens de doentes tratados no estadio 2. Considerando os 10 hospitais nesta categoria 5 são do Tipo IV, 2 dos Tipos II e V e 1 do Tipo III.

No que se refere aos doentes mais graves, a situação é a seguinte:

- Os hospitais 87, 5 e 14 são os que tratam mais doentes mais graves. Entre os 10 hospitais que tratam doentes mais graves encontram-se 6 do Tipo I, 3 do Tipo V e 1 do Tipo III;
- Com menos doentes neste estadio de severidade estão os hospitais 67, 18 e 84. Nos 10 hospitais nesta situação, 6 são do Tipo IV e 4 do Tipo II.

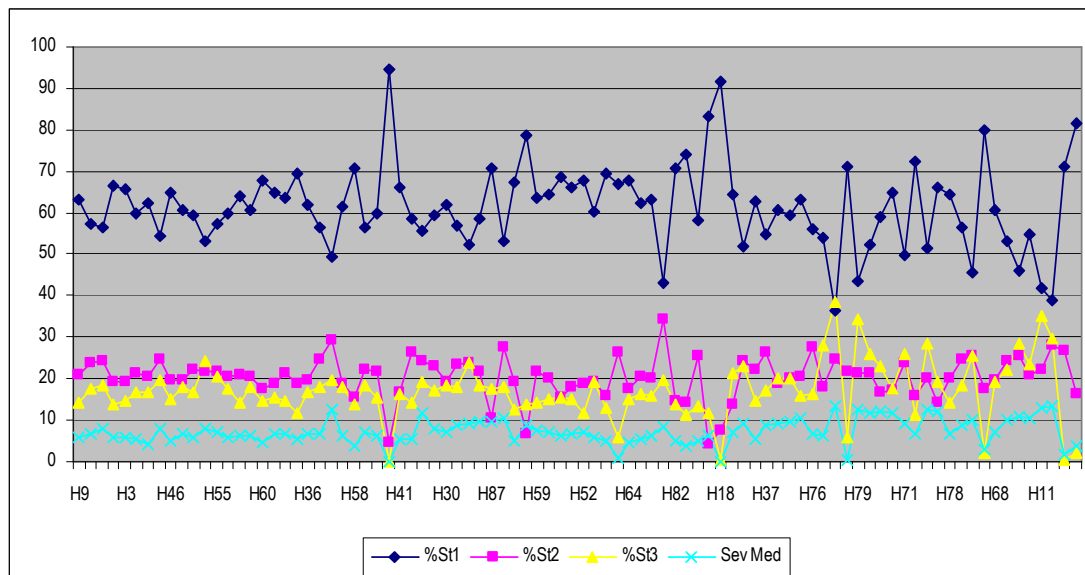
Para a severidade média, podem tecer-se os seguintes comentários:

- Os hospitais 87, 14 e 6 são os que tratam doentes com maior gravidade média. Nos 10 hospitais nesta situação, 8 são do Tipo I e 2 do Tipo V;
- Com menor gravidade média encontram-se os hospitais 18, 20 e 51. Nos 10 hospitais com menor gravidade média em todos os doentes tratados, encontram-se 5 do Tipo IV, 4 do III e 1 do III.

Nos casos cirúrgicos, à semelhança do que se observou para o total de doentes existe uma maior heterogeneidade no comportamento dos hospitais à medida em que a gravidade dos doentes aumenta, sendo os coeficientes de variação de 0.09, de 0.38 e de 1.20, respectivamente para os estadios 1, 2 e 3. Para a severidade média o coeficiente de variação é de 0.64.

Na Figura 23 são apresentadas, por hospital, as percentagens de doentes admitidos nos estadios de severidade 1 (menos grave), 2 (intermédio) e 3 (mais grave), bem como a gravidade média dos doentes, para os casos médicos. Os hospitais estão ordenados decrescentemente pela sua produção. Ver Anexo VII para se consultarem os respectivos valores.

Figura 23
% de Doentes por Estadios de Severidade e Severidade Média por Hospital
Casos Médicos



Em relação ao nível de gravidade mais baixo (estadio 1) podem fazer-se os seguintes comentários:

- Os hospitais que apresentam mais doentes neste estadio são os números 73, 11 e 81. Entre os 10 hospitais com maiores percentagens de doentes, encontram-se 5 do Tipo II, 2 dos Tipo IV e V e 1 do Tipo III;
- Os hospitais que apresentam menos doentes neste estadio são os números 14, 5 e 6. Nos 10 hospitais com percentagens mais baixas de doentes, 7 são do Tipo IV e 3 do Tipo I.

Para os doentes do estadio 2 destacam-se os seguintes aspectos:

- Os hospitais 14, 6 e 81 são os que apresentam maiores frequências relativas deste tipo de doentes. Para os 10 hospitais com maiores percentagens, 4 são do Tipo I e 2 dos Tipos II, III e IV;
- Os hospitais 86, 20 e 88 são os que apresentam menores percentagens de doentes tratados no estadio 2. Considerando os 10 hospitais nesta categoria 3 são do Tipo IV e V e 2 dos Tipos II e III.

No que se refere aos doentes mais graves, a situação é a seguinte:

- Os hospitais 73, 11 e 79 são os que tratam mais doentes mais graves. Entre os 10 hospitais que tratam doentes mais graves encontram-se 9 do Tipo IV e 1 do Tipo I;
- Com menos doentes neste estadio de severidade estão os hospitais 20, 18 e 22. Nos 10 hospitais nesta situação, 6 são do Tipo II e 2 dos Tipos III e IV.

Para a severidade média, podem tecer-se os seguintes comentários:

- Os hospitais 81, 73 e 11 são os que tratam doentes com maior gravidade média. Nos 10 hospitais nesta situação, 7 são do Tipo IV e 3 do Tipo I;
- Com menor gravidade média encontram-se os hospitais 18, 20 e 21. Nos 10 hospitais com menor gravidade média em todos os doentes tratados, encontram-se 5 do Tipo II, 3 do III e 1 do IV.

Embora a heterogeneidade do comportamento dos hospitais aumente com o nível de gravidade, existe uma menor variação nos casos médicos do que a observada para o total de doentes tratados e para os casos cirúrgicos. Os coeficientes de variação são de 0.16, de 0.25 e de 0.40, respectivamente para os estadios 1, 2 e 3. Para a severidade média o coeficiente de variação é de 0.41.

Em seguida, a severidade das admissões dos hospitais será sintetizado num indicador único. Para tal será utilizado o “z score” de acordo com a equação descrita na Metodologia.

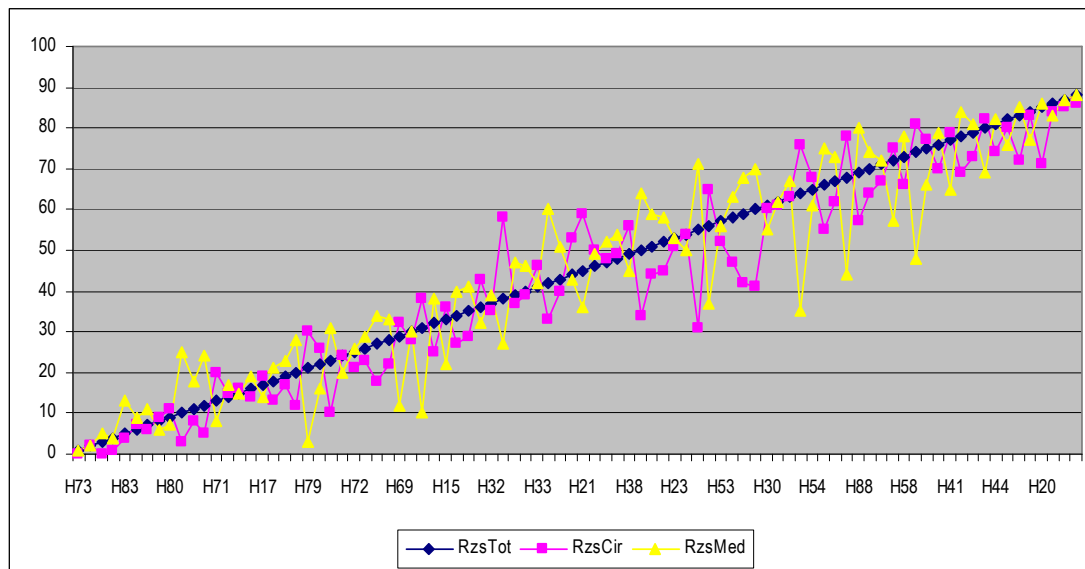
Na Figura 24 são apresentadas as ordenações dos hospitais, para o total de doentes e casos cirúrgicos e médicos.

Para o total de doentes observa-se o seguinte:

- O hospital com um “z score” menor (menos doentes do estadio 1) foi o número 73, enquanto que o hospital 9 apresenta o “z score” mais elevado (mais doentes no estadio 1);

- Entre os 10 hospitais com “z score” menor encontram-se 8 do Tipo IV e 2 do Tipo I. Entre estes hospitais 5 estão localizados na Região de Saúde III, 2 nas Regiões IV e V e 1 na Região I;
- Para os 10 hospitais com maior “z score” encontram-se 6 do Tipo I, 3 do Tipo III e 1 do Tipo II. Destes 5 estão localizados na Região V, 3 na IV e 2 na Região III.

Figura 24
Ordenação dos Hospitais por Gravidade (“Z score”)
Total de Doentes e Casos Cirúrgicos e Médicos



Para os casos cirúrgicos observa-se o seguinte:

- O hospital 81 apresenta o “z score” mais baixo, enquanto que o hospital 9 apresenta o valor mais elevado para este indicador;
- Para os 10 hospitais com o “z score” mais baixo encontram-se 9 do Tipo IV e 1 do Tipo I, estando 4 localizados na Região III, 3 na Região V, 2 na Região IV e 1 na Região I;
- Para os 10 hospitais com maior “z score” 7 são do Tipo I e 3 do Tipo III, estando 4 localizados na Região V e 3 nas Regiões III e IV;

Para os casos médicos a situação é a seguinte:

- Os hospitais 73 e 9 são os que apresentam menor e maior “z scores”, respectivamente;
- Para os 10 hospitais com menor “z score” encontram-se 7 do Tipo IV e 3 do Tipo I, estando 4 localizados na Região III, 3 na Região V, 2 na Região IV e na Região I;

- Para os 10 hospitais com maior “z score” encontram-se 5 do Tipo III, 3 do Tipo I e 1 dos Tipos II e V. Destes 5 estão localizados na Região V, 3 na Região IV e 2 na III.

Em conclusão pode afirmar-se que existem comportamentos distintos entre os casos médicos e cirúrgicos sendo o *K de Cohen* de 0.059. Isto origina igualmente que a relação entre o total de doentes e os casos cirúrgicos e os casos médicos seja igualmente reduzida, assumindo o *K de Cohen*, respectivamente valores de 0.106 e de 0.071.

Também se pode referir que não existe qualquer padrão regional para a política de admissões, visto que é frequente encontrar hospitais na mesma região que apresentam menores e maiores “z scores”.

Para o Tipo de hospital, pesem embora as diferenças apontadas para os casos médicos e cirúrgicos, pode referir-se que são os hospitais do Tipo IV os que apresentam menor proporção de doentes menos graves, enquanto que os hospitais do Tipo I são os que tratam mais doentes menos graves, para o Total de Doentes e para os casos cirúrgicos e os hospitais do Tipo III para os casos médicos.

Comparação entre complexidade e gravidade dos casos tratados

A comparação entre o desempenho dos hospitais no que respeita à sua política de admissões, caracterizada pela complexidade e gravidade dos casos tratados será feita com recurso aos “z scores”, calculados segundo a metodologia descrita anteriormente.

Para maior compreensão da análise será igualmente calculada a diferença entre o “z score” da gravidade e o “z score” da severidade.

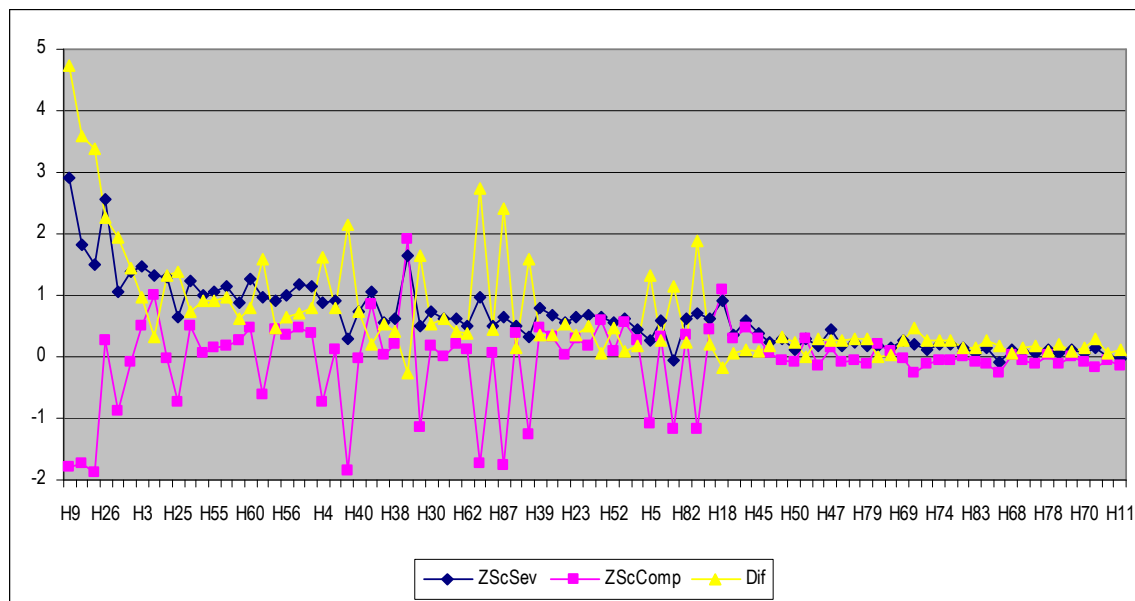
Na Figura 25 são apresentados, para todas as admissões, os “z scores” da complexidade (ZScComp) e gravidade (ZScSev) dos casos tratados, bem como a respectiva diferença (Dif), estando os hospitais ordenados decrescentemente pela sua produção. Para se consultar os valores ver Anexo VIII.

Estes elementos permitem os seguintes comentários:

- Os hospitais 9, 10 e 13 são os que apresentam maiores diferenças entre a complexidade e a gravidade dos casos tratados, sendo de referir que nestas organizações de saúde, a complexidade dos casos é sempre superior à gravidade dos mesmos;
- Para os 10 hospitais com maiores diferenças entre a complexidade e a gravidade dos casos tratados, igualmente sempre com valores superiores para o primeiro atributo, 6 são do Tipo I, 3 do Tipo V e 1 do Tipo III;

- Apenas existem 3 hospitais (20, 18 e 63, 2 do Tipo II e o restante do Tipo IV) em que a gravidade dos casos tratados é superior à complexidade;
- Os hospitais 51, 63 e 76 são os que apresentam menores diferenças entre a complexidade e a gravidade dos casos tratados;
- Para os 10 hospitais com menores diferenças entre a complexidade e a gravidade dos casos tratados 5 são do Tipo IV, 4 do Tipo III e 1 do Tipo II.

Figura 25
“Z score” das Admissões por Hospital
Complexidade e Gravidade
Todas as Admissões



Por outro lado a complexidade e gravidade dos casos por hospital apresenta uma fraca associação, visto que o coeficiente de correlação entre os “z scores” destes dois atributos é de -0.017.

No entanto, existe uma forte associação entre a quantidade produzida pelos hospitais e a gravidade dos casos tratados, sendo o respectivo coeficiente de correlação de 0.905, enquanto que esta estatística quando se considera a complexidade dos casos e o volume de produção é de -0.293, sendo ambas significativas a um nível de significância de 0.01.

Estes valores parecem traduzir que quanto maior a produção dos hospitais menor é a gravidade dos doentes tratados e que, quanto maior é a produção dos hospitais maior é a complexidade dos casos tratados.

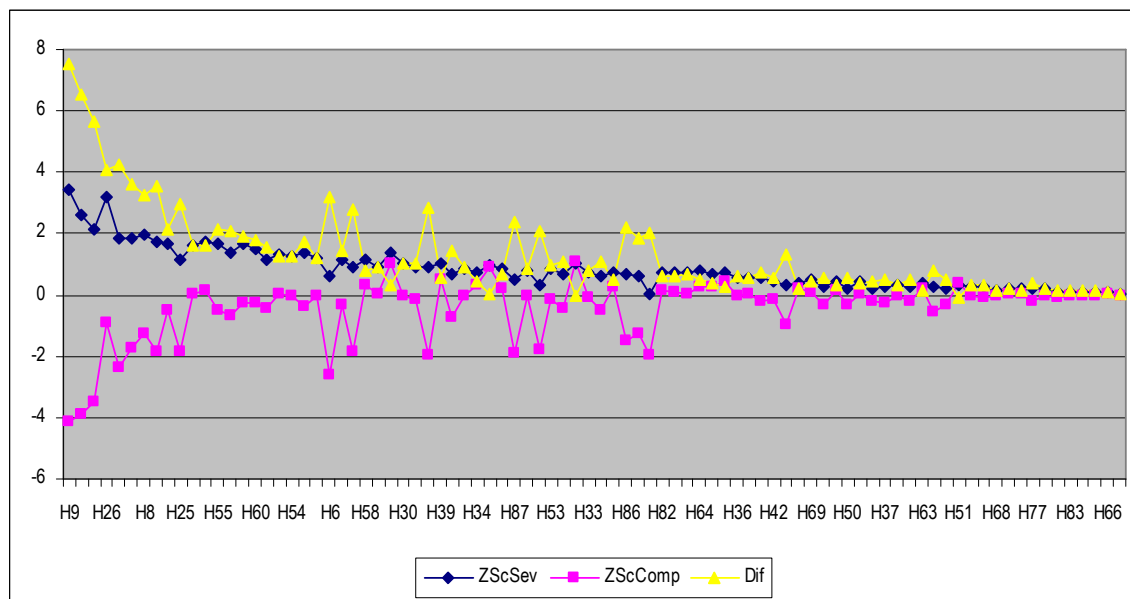
Este efeito dimensão da produção está ainda presente quando se associa esta variável com a diferença no desempenho dos hospitais entre complexidade e

gravidade dos casos tratados, visto que o coeficiente de correlação é de 0.792 (igualmente significativo a 0.01).

Neste sentido, pode concluir-se que à medida que aumenta a produção dos hospitais maior é a diferença entre a complexidade e a gravidade dos casos, pelo que poderá indiciar uma maior escolha de doentes por parte destes hospitais.

Na Figura 26 são apresentados, para os casos cirúrgicos, os “z scores” da complexidade (ZScComp) e gravidade (ZScSev) dos casos tratados, bem como a respectiva diferença (Dif), estando os hospitais ordenados decrescentemente pela sua produção. Para se consultar os valores ver Anexo VIII.

Figura 26
“Z score” das Admissões por Hospital
Complexidade e Gravidade
Casos Cirúrgicos



Podem fazer-se as seguintes observações:

- Os hospitais 9, 10 e 13 são os que apresentam maiores diferenças entre a complexidade e a gravidade dos casos tratados, sendo o primeiro atributo sempre superior à gravidade;
- Para os hospitais com maiores diferenças entre a complexidade e a gravidade dos casos, sempre com a complexidade a apresentar valores superiores, 8 são do Tipo I e 2 do Tipo III;
- Os hospitais 18, 21 e 81 são os que apresentam menores diferenças entre estes dois atributos, com a particularidade de no hospital 18 a gravidade ser superior à complexidade. Tal somente ocorre em mais um hospital (hospital 51);

- Para os 10 hospitais com menores diferenças entre a complexidade e a gravidade dos casos 7 são do Tipo IV, 2 do Tipo II e 1 do Tipo III.

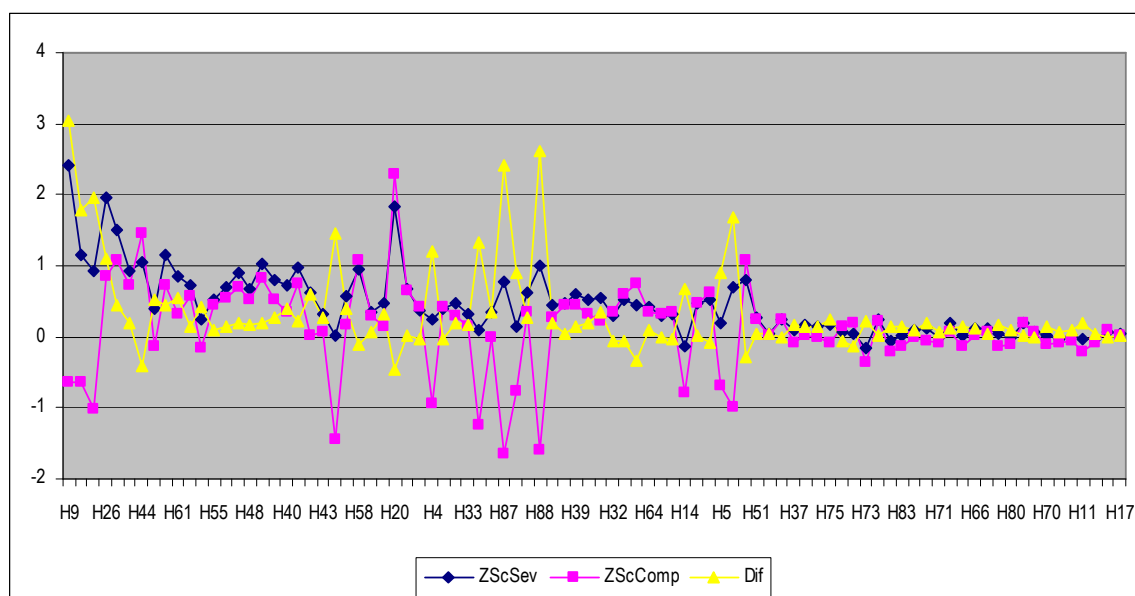
Existe igualmente uma associação entre a complexidade e a gravidade dos casos, visto que o coeficiente de correlação é de -0.476. No entanto, deve referir-se que esta associação é inversa, ou seja, quanto mais casos complexos tratados, menos casos graves são admitidos por hospital, sendo o inverso igualmente válido.

Por outro lado, existe igualmente uma associação forte entre a produção dos hospitais e a complexidade e a gravidade dos casos. Contudo, enquanto que para o primeiro atributo esta relação é directa, aumentando a complexidade com incrementos da produção (coeficiente de correlação de -0.692), com a gravidade é inversa, visto que diminui à medida que aumenta a produção (coeficiente de correlação de 0.95).

Finalmente, deve referir-se que o coeficiente de correlação entre a produção dos hospitais e a diferença entre a complexidade e a gravidade dos casos é de 0.92, o que parece igualmente indiciar um incremento na diferença do comportamento dos hospitais em relação à sua política de admissões à medida que a produção aumenta.

Na Figura 27 são apresentados, para os casos médicos, os “z scores” da complexidade (ZScComp) e gravidade (ZScSev) dos casos tratados, bem como a respectiva diferença (Dif), estando os hospitais ordenados decrescentemente pela sua produção. Para se consultar os valores ver Anexo VIII.

Figura 27
“Z score” das Admissões por Hospital
Complexidade e Gravidade
Casos Médicos



Estes elementos permitem os seguintes comentários:

- Os hospitais 9, 88 e 87 são os que apresentam maiores diferenças entre a complexidade e a gravidade dos casos, sendo o primeiro atributo sempre superior;
- Para os 10 hospitais com maiores diferenças entre a complexidade e a gravidade dos casos 6 são do Tipo I, 3 do Tipo V e 1 do Tipo III, mais uma vez com a complexidade a ser sempre superior à gravidade dos casos;
- Existem 16 hospitais em que a gravidade dos casos é superior à sua complexidade, sendo os hospitais 20, 44 e 19 onde se observam maiores diferenças;
- Os hospitais 68, 82 e 47 são os que apresentam menores diferenças entre a complexidade e a gravidade dos casos tratados, ainda assim com a complexidade a ser superior à gravidade dos casos tratados;
- Para os 10 hospitais com menores diferenças entre estes dois atributos 5 são do Tipo III, 3 do Tipo II e 2 do Tipo IV, devendo referir-se que somente 3 apresentam a gravidade superior à complexidade dos casos.

A complexidade e a gravidade dos casos apresenta uma associação directa e significativa, sendo o coeficiente de correlação de 0.37, pelo que a gravidade e a complexidade dos casos revela o mesmo comportamento, ou seja quando uma aumenta a outra também e vice-versa.

Existe igualmente uma associação inversa e significativa entre a gravidade dos casos e o volume de produção (coeficiente de correlação de 0.79), isto é quanto maior a produção maior é a proporção de casos menos graves tratados por hospital. Para a complexidade, não se detecta nenhuma associação com o volume de produção por hospital, visto que o coeficiente de correlação é de 0.07.

Finalmente, deve ainda referir-se que existe uma associação directa e significativa entre o volume da produção por hospital e a diferença do desempenho dos hospitais em função da complexidade e da gravidade dos casos tratados, sendo o coeficiente de correlação de 0.52.

Até ao presente momento, foram genericamente identificadas as seguintes situações:

- Diferenças no comportamento dos hospitais no que respeita à complexidade e à gravidade dos casos tratados, sendo de evidenciar que as diferenças entre estes dois atributos variam de forma directamente proporcional com o aumento da produção dos hospitais e que está presente para todas as admissões hospitalares, bem como para os casos cirúrgicos e médicos;
- Existência de uma relação inversa entre a complexidade e a gravidade dos casos tratados, para os casos cirúrgicos. Para os casos médicos esta relação é directa e para o total das admissões não existe qualquer

- associação entre a complexidade e gravidade dos casos tratados por hospital;
- Existência de uma relação inversa entre o volume da produção e a gravidade dos casos tratados por hospital, ou seja esta diminui à medida que aumenta a produção dos hospitais, tanto para todas as admissões, como para os casos cirúrgicos e médicos;
 - Existência de uma relação directa entre a complexidade dos casos tratados e o volume da produção por hospital, para todas as admissões e para os casos cirúrgicos. Para os casos médicos não existe nenhuma associação entre a gravidade e o volume da produção.

Neste sentido, estes resultados parecem indiciar um comportamento distinto dos hospitais no que se refere à sua política de admissões e inclusivamente que existe uma escolha de doentes por parte destas organizações de saúde, nomeadamente aquelas que apresentam maiores volumes de produção.

Para melhor se visualizar a dimensão do fenómeno observem-se os valores do Quadro XXV, onde são apresentados por quartis, a produção média dos hospitais, bem como o valor médio dos “z scores” para a complexidade e gravidade dos casos tratados e a diferença entre estes “z scores”, para todas as admissões e para os casos cirúrgicos e médicos.

Para todas as admissões a complexidade média dos casos, medida pelo “z score”, aumenta à medida que a produção dos hospitais é maior. Por sua vez a gravidade média, igualmente medida pelo “z score”, diminui à medida que a produção dos hospitais é maior, com uma única excepção encontrada entre o grupo de hospitais com menor produção e o 2º grupo de hospitais com produção menor. Finalmente, a diferença entre a complexidade dos casos e a sua gravidade aumenta proporcionalmente com o aumento da produção.

Para os casos cirúrgicos, a complexidade média dos casos, bem como a diferença entre a complexidade e a gravidade variam proporcionalmente com a produção, enquanto que a gravidade média varia inversamente com a produção dos hospitais.

Para os casos médicos, enquanto que a complexidade média dos casos e a diferença entre a complexidade e a gravidade variam proporcionalmente com a produção, a gravidade média dos casos apresenta um comportamento distinto.

Para o grupo de hospitais com menor produção a gravidade média é superior à complexidade média dos casos tratados, comportamento que se altera para os restantes três grupos de hospitais. Contudo, enquanto que entre os 1º e 2º grupos de hospitais e entre os 3º e 4º grupos de hospitais a gravidade média aumenta, entre os 2º e 3º grupos de hospitais a gravidade média diminui.

Quadro XXV
Produção Média, Complexidade Média, Gravidade Média e
Diferenças entre Complexidade e Gravidade por Hospital
por Quartis de Produção
Todas as Admissões, Casos Cirúrgicos e Casos Médicos

Todas Admissões

	Hospitais	Produção Média	Complexidade Média	Gravidade Média	Diferença Complexidade Gravidade
Até 1º Quartil	22	5206	0.110	-0.070	0.181
Entre 1 e 2º Quartis	22	14808	0.417	0.061	0.356
Entre 2 e 3º Quartis	22	27442	0.687	-0.110	0.796
Depois 3º Quartil	22	57780	1.305	-0.156	1.461

Casos Cirúrgicos

	Hospitais	Produção Média	Complexidade Média	Gravidade Média	Diferença Complexidade Gravidade
Até 1º Quartil	21	2054	0.201	-0.068	0.269
Entre 1 e 2º Quartis	22	6157	0.538	-0.233	0.771
Entre 2 e 3º Quartis	22	10182	0.864	-0.344	1.208
Depois 3º Quartil	21	23567	1.784	-1.159	2.944

Casos Médicos

	Hospitais	Produção Média	Complexidade Média	Gravidade Média	Diferença Complexidade Gravidade
Até 1º Quartil	22	3064	0.047	-0.050	0.098
Entre 1 e 2º Quartis	22	8857	0.301	0.146	0.155
Entre 2 e 3º Quartis	22	17050	0.528	0.003	0.526
Depois 3º Quartil	22	35467	0.934	0.356	0.578

Estes resultados comprovam as análises anteriormente feitas, pelo que na ausência de outras indicações parecem indiciar uma política de admissão diferente por parte dos hospitais, com os hospitais de maior dimensão a tratarem simultaneamente doentes mais complexos e menos graves e com os hospitais de menor dimensão a não apresentarem diferenças substanciais entre estes dois atributos.

Estas conclusões podem, pelo menos, ser parcialmente infirmadas se esta diferença entre a complexidade e a gravidade dos casos constituir um fenómeno estrutural, pelo que se analisará de seguida este comportamento por doença.

Para se identificarem as doenças serão utilizados os Grandes Agrupamentos de Doenças definidos pelo Disease Staging.

No Quadro XXVI são apresentadas as percentagens de doentes tratados por estadio de gravidade e por nível de complexidade, por doença.

Quadro XXVI
% Doentes Tratados por Estadio de Gravidade e por Nível de Complexidade
por Grandes Agrupamentos de Doenças (GADs)

GADs	% Estadio 1	% Estadio 2	% Estadio 3	% Nível 1	% Nível 2	% Nível 3
BR	82.86	2.93	14.08	33.88	48.35	17.77
CN	37.86	14.16	47.97	27.71	61.75	10.54
CV	22.62	41.91	35.47	15.37	62.01	22.62
DE	79.06	18.89	2.04	58.03	27.63	14.34
EN	89.79	7.46	2.74	66.86	27.36	5.78
GI	66.23	25.95	7.80	62.56	18.43	19.01
GY	86.50	11.24	2.25	36.55	52.48	10.97
HB	58.15	35.55	6.30	24.89	63.84	11.28
HE	70.93	20.15	6.27	1.53	66.35	32.12
IM	64.05	6.75	29.20	0.65	10.21	89.15
LY	87.88	7.03	5.04	0.05	34.98	64.96
MG	91.80	2.76	5.41	44.04	37.57	18.40
MS	71.15	26.87	1.98	26.42	40.26	33.33
ND	48.66	33.32	18.00	28.23	58.08	13.70
OB	94.44	5.41	0.15	98.40	1.58	0.02
OP	79.28	20.67	0.06	82.48	17.45	0.08
PN	94.68	4.30	1.02	78.26	16.81	4.93
PS	43.65	34.81	21.54	4.84	91.76	3.40
RN	60.90	19.97	19.06	42.75	40.27	16.98
RS	62.14	17.49	20.37	32.67	57.15	10.18
SY	50.02	30.26	19.72	37.99	53.02	8.98
VS	65.56	32.54	1.90	23.94	52.69	23.37
NN	58.43	0.31	0.24	90.67	6.05	3.28
OT	91.37	5.02	3.44	39.96	9.47	50.56
Total	68.50	19.31	11.48	48.52	36.25	15.23

Nota: A legenda das GADs consta do Anexo I

Para a gravidade dos casos tratados a % total dos doentes nos estadios 1, 2 e 3 é de 99.29%, dado que os restantes 0.71% correspondem a doentes a que foi atribuído o estadio 0.

Para o estadio 1, a variação entre as doenças é reduzida (coeficiente de variação de 0.28), com o máximo de admissões a ser atingido nas Doenças do sistema nervoso periférico e o mínimo nas Doenças cardiovasculares.

As situações em que se observa uma maior percentagem de doentes são as Doenças do sistema nervoso periférico, os Casos obstétricos, as Doenças dos órgãos genitais masculinos, as Outras doenças e as Doenças dos olhos, do nariz e da garganta.

Na situação contrária, ou seja, com uma menor percentagem de doentes encontram-se as Doenças cardiovasculares, as Doenças do sistema nervoso central, as Doenças psiquiátricas, as Doenças endócrinas e metabólicas e as Doenças multisistémicas.

Para o estadio 2 a heterogeneidade entre doenças aumenta, sendo o coeficiente de variação de 0.71, com um máximo nas Doenças cardiovasculares e um mínimo das Doenças não especificadas.

Com mais doentes admitidos encontram-se as Doenças cardiovasculares, as Doenças hepatobiliares, as Doenças psiquiátricas, as Doenças endócrinas e metabólicas e as Doenças vasculares (excepto as cardíacas).

Com menos doentes admitidos encontram-se as Doenças não especificadas, as Doenças dos órgãos genitais masculinos, as Doenças da mama, as Doenças do sistema nervoso periférico e as Outras doenças.

No estadio 3 as diferenças entre as doenças são ainda mais relevantes, visto que o coeficiente de variação é de 1.12, com um máximo nas Doenças do sistema nervoso central e um mínimo nas Doenças oftalmológicas.

Com mais doentes admitidos encontram-se as Doenças do sistema nervoso central, as Doenças cardiovasculares, as Doenças imunológicas, as Doenças psiquiátricas e as Doenças respiratórias.

Com menos doentes admitidos encontram-se as Doenças oftalmológicas, os Casos obstétricos, as Doenças não especificadas, as Doenças do sistema nervoso periférico e as Doenças vasculares (excepto as cardíacas).

Para o Nível 1 de complexidade existe alguma variação entre as doenças (coeficiente de variação de 0.71), com o máximo a ser atingido nos Casos obstétricos e o mínimo nas Doenças linfáticas.

As doenças onde se observa uma maior percentagem de admissões são os Casos obstétricos, as Doenças não especificadas, as Doenças oftalmológicas, as Doenças do sistema nervoso periférico e as Doenças dos olhos, do nariz e da garganta.

Com uma menor percentagem de admissões encontram-se as Doenças linfáticas, as Doenças imunológicas, as Doenças hematológicas, as Doenças psiquiátricas e as Doenças cardiovasculares.

Para o Nível 2 de complexidade a variação entre doenças, embora ainda seja relevante é menor do que no nível 1 (coeficiente de variação de 0.58), com o máximo a ser atingido nas Doenças psiquiátricas e o mínimo nos Casos obstétricos.

Com mais admissões encontram-se as Doenças psiquiátricas, as Doenças hematológicas, as Doenças hepatobiliares, as Doenças cardiovasculares e as Doenças do sistema nervoso central.

Com menos admissões encontram-se os Casos obstétricos, as Doenças não especificadas, as Outras doenças, as Doenças imunológicas e as Doenças do sistema nervoso periférico.

Para o Nível 3 de complexidade observa-se a maior variação entre doenças (coeficiente de variação de 1.05), com o máximo nas Doenças imunológicas e o mínimo nos Casos obstétricos.

Com mais admissões encontram-se as Doenças imunológicas, as Doenças linfáticas, as Outras doenças, as Doenças do sistema musculoesquelético e as Doenças hematológicas.

Com menos admissões encontram-se os Casos obstétricos, as Doenças oftalmológicas, as Doenças não especificadas, as Doenças psiquiátricas e as Doenças do sistema nervoso periférico.

Para se pesquisar a existência de comportamentos distintos por doença em função da complexidade e da gravidade dos casos tratados serão construídos “z scores” de acordo com as equações apresentadas na Metodologia.

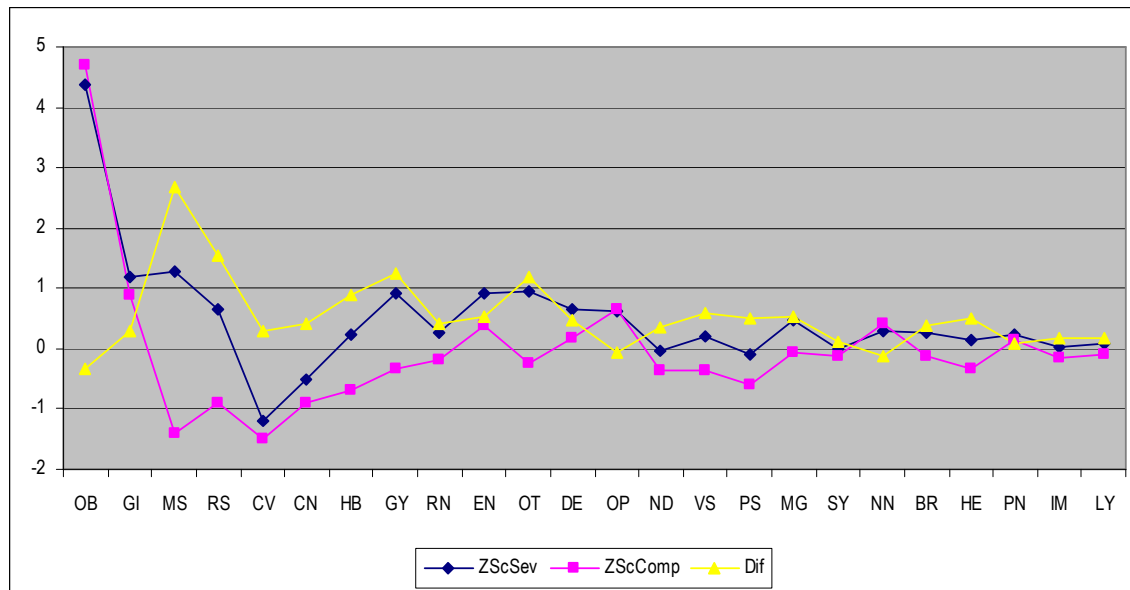
Na Figura 28 são apresentados, para todas as admissões, os “z scores” da complexidade (ZScComp) e gravidade (ZscSev) dos casos tratados, bem como a respectiva diferença (Dif), estando as doenças ordenadas decrescentemente pela sua produção. Para se consultar os valores ver Anexo IX.

As doenças com casos mais complexos são as cardiovasculares, do sistema musculoesquelético, do sistema nervoso central, as respiratórias e as Doenças hepatobiliares, enquanto que as que apresentam casos menos complexos são os Casos obstétricos, as Doenças gastrointestinais, as Doenças oftalmológicas, as Doenças não especificadas e as Doenças dos olhos, do nariz e da garganta.

Por sua vez, as doenças com casos mais graves são as cardiovasculares, do sistema nervoso central, as psiquiátricas, as endócrinas e metabólicas e as multisistêmicas. Com casos menos graves encontram-se os Casos obstétricos, as Doenças do sistema musculoesquelético, as Doenças gastrointestinais, as Outras doenças e as Doenças dos olhos, do nariz e da garganta.

Finalmente, as Doenças do sistema musculoesquelético, as Doenças respiratórias, as Doenças ginecológicas, as Outras doenças e as Doenças hepatobiliares são aquelas onde se encontram maiores diferenças entre a complexidade e a gravidade dos casos.

Figura 28
“Z score” das Admissões por Doença,
Complexidade e Gravidade
Todas as Admissões



No entanto, apesar das diferenças apontadas a associação entre a complexidade e a gravidade dos casos por doença (ambas medidas pelo “z score”) é bastante forte e positiva (coeficiente de correlação de 0.84).

Atendendo a que esta associação é igualmente encontrada para os casos cirúrgicos e médicos, sendo os coeficientes de correlação respectivamente de 0.25 e de 0,87 (ver Anexo IX), pode concluir-se que, por doença, quanto mais casos tratados com maior complexidade, mais casos graves são admitidos.

Neste sentido, pode concluir-se que não existe nenhuma razão estrutural, sendo esta aferida pelo tipo de doença, para existirem associações inversas entre a complexidade e a gravidade dos casos tratados, pelo que se reitera o referido anteriormente, essencialmente que os hospitais com maior dimensão tratam doentes com maior complexidade e com menor gravidade.

Esta situação parece configurar um cenário típico de diferentes perfis das admissões por parte dos hospitais, sendo as maiores diferenças entre a complexidade e a severidade encontradas nos hospitais com maior dimensão.

Assim, em seguida serão analisados os potenciais benefícios (resultados brutos) para os hospitais resultantes dos diferentes perfis de admissão atrás descritos.

Os resultados brutos dos hospitais serão medidos pela demora média e pela taxa de mortalidade.

Para se caracterizar o perfil das admissões dos hospitais duas alternativas se colocam:

- Utilizar os “z scores” da complexidade e da gravidade dos casos admitidos por hospital;
- Utilizar os Índices de Casemix do Disease Staging e dos DRGs.

Em termos teóricos os Índices de Casemix capturam melhor a dimensão do perfil das admissões, não só pelo facto de utilizarem uma escala mais poderosa, mas também porque qualificam melhor as diferenças da gravidade, inclusivamente dentro do mesmo estadio, para as diferentes doenças.

No Quadro XXVII são apresentados os coeficientes de correlação entre o perfil das admissões dos hospitais medidos pelo “z score” e pelos Índices de Casemix.

Quadro XXVII
Coeficientes de Correlação; Perfil das Admissões Hospitalares

	Índice Casemix Disease Staging	Índice Casemix DRGs	Índice Casemix Composto
Gravidade	-0.408 *	-0.046	-0.334 *
Complexidade	-0.408 *	-0.711 *	-0.638 *
Composto	-0.574 *	-0.639 *	-0.711 *

* significativo a 0.01

Como se verifica para todos os apuramentos os dois indicadores estão a dar a mesma informação, visto que todas as correlações são significativas e exprimem a relação esperada. A única excepção é encontrada na associação entre a gravidade medida pelo “z score” e a complexidade medida pelo Índice de Casemix dos DRGs.

Contudo, para os aspectos mais relevantes nesta análise os dois indicadores estão a dar indicações semelhantes. Ou seja, quanto maior o número de doentes graves admitidos, maior a sua gravidade média e quanto mais casos complexos tratados maior é a complexidade média.

Assim, pode afirmar-se que a utilização dos Índices de Casemix do Disease Staging, dos DRGs e Composto podem ser utilizados como indicadores de síntese do perfil das admissões dos hospitais (carácter analítico).

No entanto, para melhor se descrever o perfil das admissões, tanto para a totalidade dos episódios de internamento, como para os casos cirúrgicos e para os casos médicos e ainda por doença, os níveis de complexidade e os estadios de gravidade e, naturalmente os “z scores”, podem revelar maior utilidade (carácter descritivo).

Finalmente, será calculada a diferença entre a gravidade e a complexidade dos casos tratados de acordo com a seguinte equação:

Diferença entre Gravidade e Complexidade dos casos admitidos = Índice de Casemix do Disease Staging – Índice de Casemix dos DRGs

Este indicador (diferença entre a gravidade e a complexidade) é positivo quando a gravidade é superior à complexidade e negativo na situação inversa. Assim, deve ter-se presente que um valor positivo na associação com os resultados brutos traduz, por exemplo, que a mortalidade diminui à medida que existem mais casos complexos.

No Quadro XXVIII são apresentados os coeficientes de correlação entre o perfil das admissões e os respectivos resultados brutos.

Quadro XXVIII
Coeficientes de Correlação por Hospital entre
Perfil das Admissões e Resultados Brutos

	Taxa de Mortalidade	Demora Média
Complexidade	0.250 *	0.645 **
Gravidade	0.945 **	0.687 **
Composto	0.844 **	0.803 **
Diferença	0.821 **	0.319 **

* significativo a 0.05

** significativo a 0.01

Para a complexidade e para a gravidade dos casos tratados, tanto consideradas isoladamente, como conjuntamente, todas as associações são significativas.

Como principais observações podem referir-se as seguintes:

- A taxa de mortalidade diminui à medida que os hospitais tratam doentes menos graves, menos complexos e menos complexos e graves;
- A demora média diminui à medida que os hospitais tratam doentes menos graves, menos complexos e menos complexos e graves;
- A demora média e a taxa de mortalidade diminuem à medida em que a diferença entre a complexidade e a gravidade dos casos é maior (mais casos complexos e menos graves).

Todos estes aspectos permitem desde já evidenciar que, na ausência de um sistema de medição da gravidade e de ajustamento pelo risco, os hospitais são duplamente beneficiados quando optimizam a diferença entre a complexidade e a gravidade dos casos.

Em primeiro lugar, porque os seus resultados brutos, medidos pela demora média e pela taxa de mortalidade variam de uma forma inversa com a gravidade dos doentes.

Em segundo lugar, porque a neutralidade do sistema de financiamento em relação à gravidade dos doentes, medida pelo respectivo Índice de Casemix, pode trazer vantagens para os hospitais que tratam doentes menos graves, desde que se assuma que existe uma relação directa entre gravidade dos doentes e utilização de recursos.

5.2. Desempenho

5.2.1. Efectividade

Conforme foi referido a avaliação da efectividade dos hospitais será feita tendo em conta um indicador de resultados finais – a mortalidade.

De acordo com as indicações defendidas, entre outros, por Iezzoni e outros (1997c) e O’Muircheartaigh, Murphy e Moore (2002) serão utilizados dados referentes a três anos (1999 a 2001) para avaliar o desempenho dos hospitais.

É ainda de referir que as taxas de mortalidade dizem somente respeito a óbitos que ocorrem dentro do hospital, o que para além de suscitar algumas reservas quando se procede a comparações entre hospitais, essencialmente por razões atribuíveis a eventuais diferenças na política de altas dos hospitais (Iezzoni et al, 1995a), constitui, no entanto, o único elemento de análise possível na actual realidade portuguesa, visto que as informações de rotina das organizações de saúde, não permitem uma análise da mortalidade 30 dias após o episódio hospitalar.

Embora tendo presente os argumentos referidos por Fink, Yano e Brook (1989), essencialmente no que se refere à existência de diferenças na taxa de mortalidade entre diversos problemas de saúde ou doenças, o que pode limitar a importância do seu valor global, tem-se igualmente em conta os aspectos defendidos pela “Commission on Professional and Hospital Activities” (CPHA) (1987) (citado em Fink, Yano e Brook, 1989) “as taxas de mortalidade são actualmente reconhecidas como uma medida de qualidade dos cuidados prestados”.

Na mesma linha encontra-se o estudo de Dubois e outros (1987), na qual foi estabelecida uma relação directa entre as taxas de mortalidade e a qualidade dos cuidados prestados, ou ainda o defendido por Rosen e Green (1987) – “as altas taxas de mortalidade não são um problema em si mesmo, devendo a principal preocupação residir na proporção de mortes evitáveis” – e por Griffith, Alexander e Jelinek (2002), no qual é afirmado que os hospitais que apresentam melhores resultados, tanto na taxa de mortalidade ajustada pelo risco, com na taxa de complicações ajustada pelo risco, parecem demonstrar um melhor desempenho em termos de uma maior flexibilidade e de uma melhor adaptação às preferências dos consumidores.

Como foi referido, tendo igualmente presente que eventuais diferenças nas taxas de mortalidade hospitalar podem traduzir diferenças nas características dos doentes, nomeadamente as respeitantes à doença e respectiva complexidade, bem como na severidade do estado do doente, irá ser utilizado um sistema de ajustamento pelo risco – Disease Staging – para desta forma se identificarem e medirem as principais diferenças nas taxas de mortalidade dos hospitais, em função dos valores observados e dos valores previstos pelo referido sistema de classificação de doentes.

Assim, o presente capítulo é constituído por três partes:

- Caracterização da taxa de mortalidade hospitalar observada;
- Caracterização da taxa de mortalidade hospitalar esperada;
- Avaliação da Efectividade dos hospitais (taxa de mortalidade hospitalar observada versus taxa de mortalidade hospitalar esperada).

5.2.1.1. Caracterização da Taxa de Mortalidade Hospitalar Observada

Os dados em análise dizem respeito a três anos (1999, 2000 e 2001), podendo no Quadro XXIX observar-se o total de casos e as taxas de mortalidade por anos e para a globalidade do período.

Quadro XXIX
Total de Episódios e Taxas de Mortalidade Hospitalar Observadas
1999, 2000, 2001 e Total do Período

Ano	Total de Episódios	Taxa de Mortalidade (%)
1999	775390	4.58
2000	757211	4.67
2001	782588	4.69
Período	2315189	4.65

Fonte: Direcção Geral da Saúde

O número total de episódios de internamento oscila entre os 782588 em 2001 e 757211 em 2000, correspondendo a um total de 2351189 para o período em análise.

A taxa de mortalidade apresenta uma tendência moderadamente crescente, sendo de 4.65% para a média dos três anos em estudo.

No Quadro XXX apresentam-se os valores da taxa de mortalidade por Região de Saúde e por estadio de gravidade.

Quadro XXX
Taxas de Mortalidade Hospitalar Observadas por Região de Saúde e
por Estádios de Gravidade

Região	Taxa de Mortalidade Total (%)	Taxa de Mortalidade Estadio 1 (%)	Taxa de Mortalidade Estadio2 (%)	Taxa de Mortalidade Estádio 3 (%)
I	6.02	3.65	5.45	21.01
I	5.51	3.58	5.08	18.86
II	4.07	1.94	3.51	17.79
IV	5.51	2.81	4.85	22.13
V	3.95	1.94	3.78	17.29
Total	4.65	2.35	4.19	19.36

Fonte: Direcção Geral da Saúde

As Regiões de Saúde com taxas de mortalidade observada mais altas, são as I, II e IV, enquanto que a Região V apresenta a taxa de mortalidade mais baixa.

A taxa de mortalidade aumenta com a gravidade dos doentes, sendo em média cerca de 8.2 vezes superior no estadio 3 em relação ao estadio 1. Nas Regiões de Saúde I, II e IV esta razão é inferior, enquanto que nas restantes regiões se observa um comportamento contrário, com maior expressão na Região III.

Nos estadios 1 e 2 a Região I apresenta as taxas de mortalidade mais elevadas, enquanto que no estadio 3 tal ocorre na Região IV. Por outro lado, nos estadios 1 e 2 a Região III apresenta as taxas de mortalidade mais baixas, sendo a Região IV que apresenta este comportamento no estadio 3.

Analizando os dados por tipo de hospital e estadios de gravidade, observem-se os valores do Quadro XXXI.

Quadro XXXI
Taxas de Mortalidade Hospitalar Observadas por Tipo de Hospital e por Estadios de Gravidade

Tipo de Hospital	Taxa de Mortalidade Total (%)	Taxa de Mortalidade Estadio 1 (%)	Taxa de Mortalidade Estadio 2 (%)	Taxa de Mortalidade Estadio 3 (%)
I	4.96	2.41	4.25	19.90
II	0.18	0.09	0.27	5.83
III	4.53	2.33	4.18	18.81
IV	5.64	3.59	4.77	17.86
V	6.52	2.80	7.71	26.51
Total	4.65	2.35	4.19	19.36

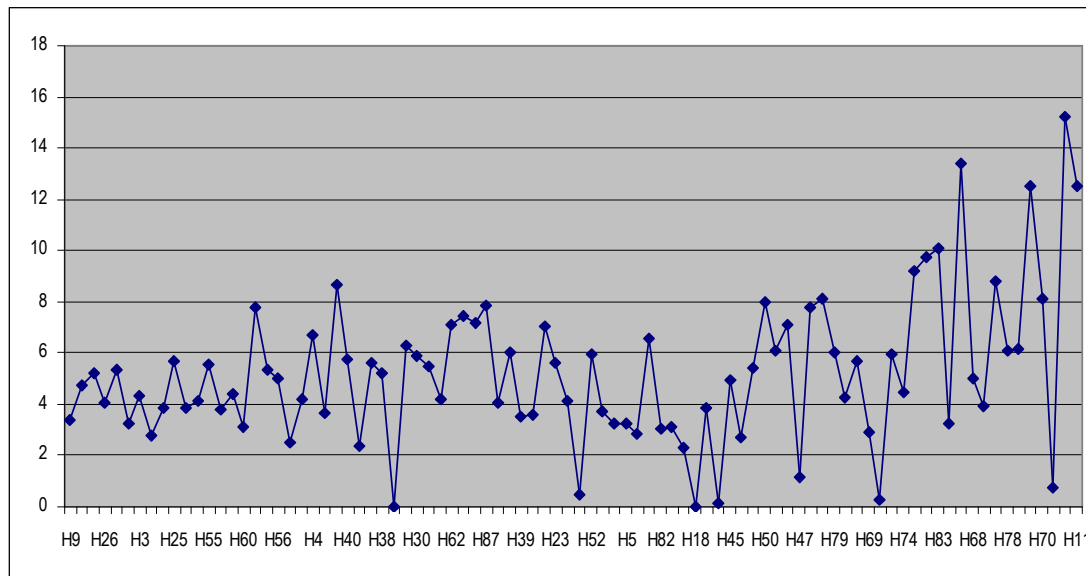
Fonte: Direcção Geral da Saúde

Estes elementos permitem evidenciar os hospitais do Tipo IV e V como aqueles que apresentam as taxas de mortalidade mais elevadas, enquanto que os hospitais do Tipo II apresentam taxas de mortalidade praticamente insignificantes.

No estadio 1 os hospitais do Tipo IV apresentam as taxas de mortalidade mais elevadas e os hospitais do Tipo V nos estadios 2 e 3. Para as taxas de mortalidade mais baixas, à semelhança do observado para as taxas de mortalidade globais, são sempre os hospitais do Tipo II que apresentam os menores valores por estadios de gravidade.

Na Figura 29 são apresentadas as taxas de mortalidade por hospital, ordenados decrescentemente pela quantidade produzida (no Anexo X são apresentadas as taxas de mortalidade por hospital).

Figura 29
Taxas de Mortalidade Observadas por Hospital



As taxas de mortalidade hospitalar variam entre 0.01% no hospital 18 e 15.25% no hospital 81, sendo ainda de evidenciar que, mesmo assim, o comportamento das taxas de mortalidade entre os 88 hospitais é relativamente homogêneo – coeficiente de variação de 0.55.

Estes valores parecem igualmente traduzir uma associação inversa entre o volume da produção e a taxa de mortalidade hospitalar, visto que o coeficiente de correlação entre estas duas variáveis é de -0.227 (significativo a 0.05), o que permite afirmar que a taxa de mortalidade observada aumenta à medida que a quantidade produzida pelos hospitais diminui.

Nos Quadros XXXII e XXXIII são identificados, respectivamente, os 10 hospitais com taxas de mortalidade mais baixas e aqueles que apresentam as taxas de mortalidade mais elevadas.

Quadro XXXII
Hospitais com Taxas de Mortalidade Observadas Mais Baixas

Hospitais	Região	Tipo de Hospital
18	V	II
20	IV	II
21	V	II
22	IV	II
19	IV	II
17	IV	II
47	III	III
85	V	IV
58	V	III
41	III	III

Em relação aos hospitais com taxas de mortalidade mais baixas observa-se uma grande predominância dos hospitais do Tipo II, estando ainda a grande maioria dos hospitais situados nas Regiões de Saúde IV e V.

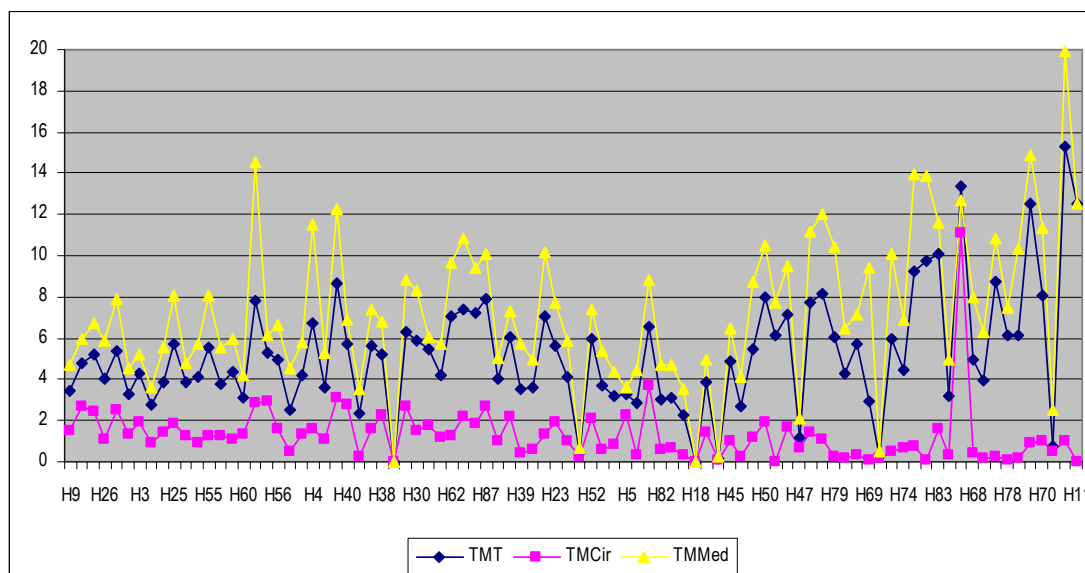
Quadro XXXIII
Hospitais com Taxas de Mortalidade Observadas Mais Elevadas

Hospitais	Região	Tipo de Hospital
75	IV	IV
6	IV	I
65	III	IV
72	II	IV
67	III	IV
83	III	IV
11	V	I
66	IV	IV
73	V	IV
81	I	IV

Em relação aos hospitais com taxas de mortalidade mais elevadas observa-se uma grande concentração nos de Tipo IV, embora com uma distribuição geográfica bem mais equilibrada.

Na Figura 30 são apresentadas as taxas de mortalidade por hospital, para o total de admissões e para os casos cirúrgicos e médicos, estando os hospitais ordenados por ordem crescente da taxa de mortalidade global (ver valores no Anexo X).

Figura 30
Taxas de Mortalidade Observadas por Hospital
Total de Admissões (TMT), Casos Cirúrgicos (TMCir)
e Casos Médicos (TMMed)



Para os casos cirúrgicos as taxas de mortalidade variam entre 0,00% (hospital 20) e 3.69% no hospital 14, ainda assim com uma relativa heterogeneidade entre hospitais, visto que o coeficiente de variação é de 0.72.

Ainda nos casos cirúrgicos e ao contrário do que se verifica para todas as admissões, existe uma associação directa entre volume de produção e a taxa de mortalidade por hospital, sendo o coeficiente de correlação de 0.399 (significativo a 0.001).

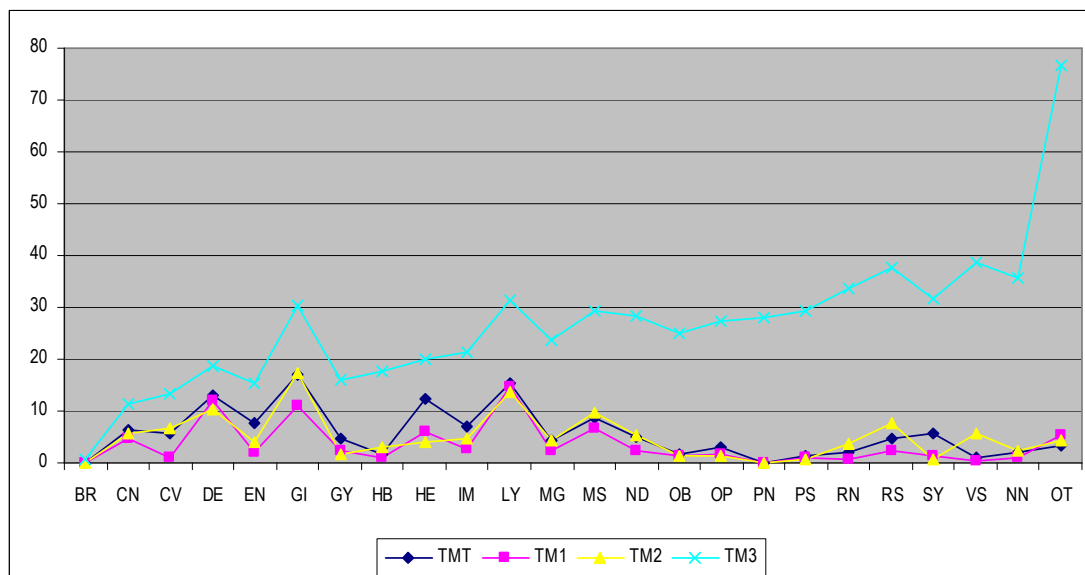
Nos casos médicos as taxas de mortalidade variam entre 0.01 no hospital 18 e 19.68% no hospital 81. No entanto, atendendo ao valor do coeficiente de variação (0.49), pode referir-se que são as situações que apresentam um comportamento menos heterogéneo entre hospitais.

Mais uma vez parece existir uma relação inversa entre o volume da produção e a taxa de mortalidade por hospital, visto que o coeficiente de correlação é de - 0.245 (significativo a 0.05).

Deve ainda referir-se que, com a excepção do hospital 18, a taxa de mortalidade nos casos médicos é sempre superior à dos casos cirúrgicos, com o valor máximo a ser encontrado no hospital 81 e o mínimo no hospital 20.

Na Figura 31 são apresentadas as taxas de mortalidade por doença, bem como pelos estadios de gravidade (no Anexo X podem ser consultados os valores). Para este efeito as doenças foram agregadas por Grandes Agrupamentos de Doenças do Disease Staging (ver Anexo I) (Gonnella, Louis e Gozum, 1999).

Figura 31
 Taxas de Mortalidade Hospitalar Observadas
 por Doença e por Estadios de Gravidade



Para todas as admissões aparecem com as taxas de mortalidade mais elevadas (TMT) as Doenças imunológicas, do sistema linfático, do aparelho respiratório, do sistema nervoso central e hematológicas.

Com taxas de mortalidade mais baixas encontram-se os Casos de obstetrícia e as Doenças oftalmológicas, do sistema nervoso periférico, as ginecológicas e as dos olhos, do nariz e da garganta.

No estadio 1 as doenças com taxas de mortalidade (TM1) mais elevadas são as Doenças do sistema linfático, respiratórias, imunológicas, hematológicas e do sistema nervoso central. Com taxas de mortalidade mais baixas encontram-se os Casos obstétricos e as Doenças oftalmológicas, do sistema nervoso periférico, Outras doenças e Doenças ginecológicas.

No estadio 2 as Doenças imunológicas, do sistema linfático, respiratórias, hematológicas e vasculares apresentam as taxas de mortalidade (TM2) mais elevadas, enquanto que os Casos obstétricos e as Doenças oftalmológicas, da mama, ginecológicas e do sistema musculoesquelético apresentam as taxas de mortalidade mais baixas.

No estadio 3 as Doenças não especificadas, do sistema nervoso periférico, vasculares, dermatológicas e Outras doenças apresentam as taxas de mortalidade (TM3) mais altas, enquanto que os Casos obstétricos e as Doenças endócrinas e metabólicas, psiquiátricas, cardiovasculares e renais apresentam as taxas de mortalidade mais baixas.

Na generalidade, as taxas de mortalidade no estadio 2 são superiores às do estadio 1, com maior expressão nas Doenças imunológicas, psiquiátricas e vasculares. Mesmo assim, existem 8 doenças (entre 24) nas quais se observam taxas de mortalidade mais elevadas no estadio 1, com maior relevo nas Doenças do sistema nervoso central, respiratórias e do sistema linfático.

Entre os estadios 3 e 2 observa-se sempre a existência de taxas de mortalidade mais elevadas nos doentes mais graves. As maiores diferenças são encontradas nas Doenças não específicas, dermatológicas e do sistema nervoso periférico, enquanto que as diferenças mais reduzidas são encontradas nos Casos obstétricos e nas Doenças endócrinas e metabólicas e psiquiátricas.

5.2.1.2. Caracterização da Taxa de Mortalidade Esperada

Tendo em atenção os aspectos metodológicos referidos no respectivo capítulo apresentar-se-á a Taxa de Mortalidade Esperada por Região, por Tipo de hospital e por Hospital.

No Quadro XXXIV são apresentadas as taxas de mortalidade esperadas por Região de Saúde.

Quadro XXXIV
Taxas de Mortalidade Hospitalar Esperadas por Região de Saúde

Região	Taxa de Mortalidade (%)
I	5.35
II	5.01
III	4.45
IV	5.25
V	4.05
Total	4.65

Os elementos do Quadro XXXIV permitem afirmar que:

- Não existe uma grande variação nas taxas de mortalidade esperadas por Região de Saúde (o coeficiente de variação é de 0.12);
- O valor máximo é encontrado na Região I e o mínimo na Região V;
- As Regiões que, em média, tratam doentes mais graves, são respectivamente a I e a IV. Na situação contrária são as Regiões V e III que tratam doentes com menor gravidade.

No Quadro XXXV são apresentadas as taxas de mortalidade esperadas por Tipo de hospital.

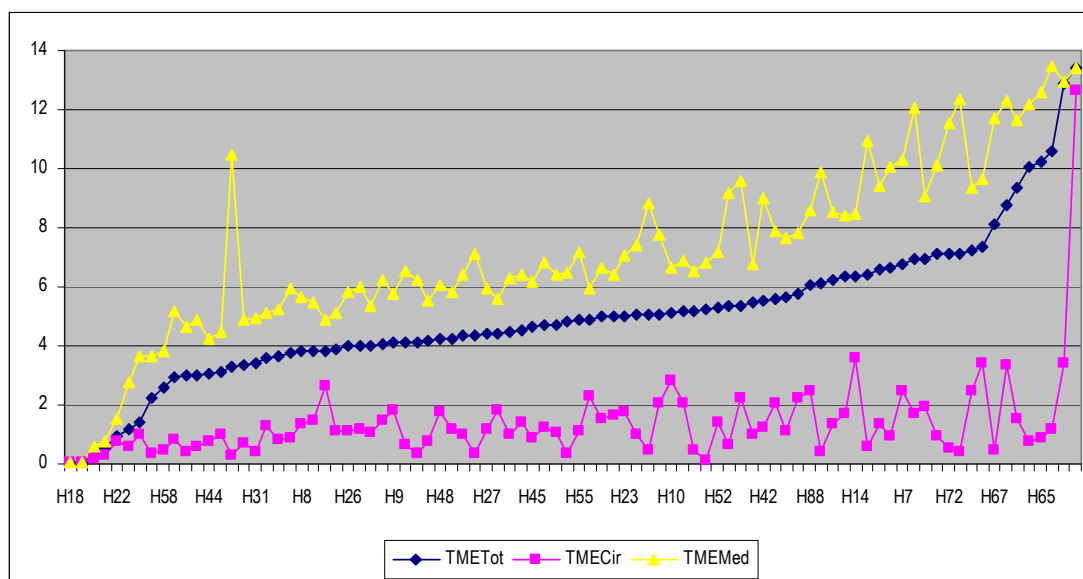
Quadro XXXV
Taxas de Mortalidade Hospitalar Esperadas por Tipo de Hospital

Tipo de Hospital	Taxa de Mortalidade (%)
I	5.19
II	0.32
III	4.41
IV	5.43
V	6.11
Total	4.65

Por Tipo de hospital existe uma maior heterogeneidade nas taxas de mortalidade esperadas, variando entre 0.32% e 6.11% e com um coeficiente de variação de 0.54. Os hospitais que tratam, em média, doentes mais graves são os do Tipo V e IV, estando da situação oposta os hospitais do Tipo II e III.

Na Figura 32 são apresentadas as Taxas de Mortalidade Esperadas por hospital, para o total das admissões, casos médicos e cirúrgicos, estando os hospitais ordenados por ordem crescente da taxa de mortalidade para todas as admissões (no Anexo X são apresentadas as taxas de mortalidade esperadas por hospital).

Figura 32
Taxas de Mortalidade Esperadas por Hospital
Todas as Admissões, Casos Cirúrgicos e Casos Médicos



Para todas as admissões, as taxas de mortalidade esperadas (TME_{Tot}) variam entre 0.03% no hospital 18 e 13.43% no hospital 73. O comportamento das taxas de mortalidade esperadas entre os 88 hospitais apresenta menor variação que as taxas de mortalidade observadas, sendo os coeficientes de variação de, respectivamente 0.48 e 0.55.

Para os casos cirúrgicos a taxa de mortalidade esperada (TME_{Cir}) por hospital varia entre 0.03% no hospital 18 e 12.66% no hospital 73. As taxas de mortalidade observadas e esperadas apresentam praticamente o mesmo comportamento por hospital, visto que os coeficientes de variação são, respectivamente de 1.06 e 1.07.

Nos casos médicos a taxa de mortalidade esperada (TME_{Med}) por hospital varia entre 0.03% no hospital 18 e 13,50% no hospital 81. Mais uma vez a homogeneidade da taxa de mortalidade esperada por hospital é ligeiramente maior que a da taxa de mortalidade observada, sendo os coeficientes de variação de 0.42 e de 0.49, respectivamente.

Nos Quadros XXXVI e XXXVII são identificados, respectivamente, os hospitais com as taxas de mortalidade observadas e esperadas mais baixas e aqueles que apresentam taxas de mortalidade mais elevadas.

Quadro XXXVI
Hospitais com Taxas de Mortalidade Observadas e Esperadas Mais Baixas, por Região e Tipo de Hospital

Taxa de Mortalidade Observada			Taxa de Mortalidade Esperada		
Hospitais	Região	Tipo Hosp.	Hospitais	Região	Tipo Hosp.
18	V	II	18	V	II
20	IV	II	20	IV	II
21	IV	II	21	V	II
22	IV	II	19	IV	II
19	IV	II	22	IV	II
17	IV	II	47	III	III
47	III	III	17	IV	II
85	V	IV	85	V	IV
58	V	III	58	V	III
41	III	III	41	III	III

Em relação aos hospitais com taxas de mortalidade esperadas mais baixas, observa-se que todos os hospitais constam da lista daqueles que apresentam uma taxa de mortalidade observada mais baixa. Nesta comparação somente se observam trocas de posições entre os hospitais 19 e 22 e entre os hospitais 47 e 17.

Para as taxas de mortalidade esperadas observa-se ainda uma grande predominância dos hospitais do Tipo II (6 hospitais), sendo os restantes do Tipo III (3 hospitais) e do Tipo IV. Por Região de Saúde 5 hospitais estão localizados na Região IV, 3 hospitais na Região V e os restantes hospitais na Região III.

Quadro XXXVII
Hospitais com Taxas de Mortalidade Observadas e Esperadas Mais Elevadas, por Região e Tipo de Hospital

Taxa de Mortalidade Observada			Taxa de Mortalidade Esperada		
Hospitais	Região	Tipo Hosp.	Hospitais	Região	Tipo Hosp.
81	I	IV	73	V	IV
73	V	IV	11	V	I
66	IV	IV	81	I	IV
11	V	I	65	III	IV
83	III	IV	66	IV	IV
67	III	IV	83	III	IV
72	II	IV	6	IV	IV
65	III	IV	67	III	IV
6	IV	I	87	IV	V
75	IV	IV	15	IV	I

Para os hospitais com taxas de mortalidade esperadas mais elevadas a situação é diferente, tanto em termos de hospitais presentes na lista, aparecendo os nºs 87 e 15 e desaparecendo os nºs 72 e 75, como em termos de posição, como, por exemplo, comprovam os comportamentos dos hospitais 81 e 67.

Contudo, em termos gerais, pode dizer-se que existe uma forte associação entre a taxa de mortalidade observada e a taxa de mortalidade esperada, visto que o coeficiente de correlação é de 0.945.

Para as taxas de mortalidade esperadas por Tipo de hospital observa-se uma grande predominância dos hospitais do Tipo IV (7 hospitais), sendo os restantes dos Tipos I (2 hospitais) e V. Por Região de Saúde observa-se uma maior dispersão geográfica, com 4 hospitais localizados na Região IV, 3 na Região III, 2 hospitais na Região V e 1 na Região II.

Tendo em atenção estes aspectos interessa agora realizar a comparação entre as taxas de mortalidade observadas e as esperadas e desta forma avaliar a efectividade dos hospitais.

5.2.1.3. Avaliação da Efectividade

A efectividade dos hospitais será medida, num primeiro momento, pela Taxa de Mortalidade Padronizada (TMP) que consiste na divisão entre os valores observados e esperados das Taxas de Mortalidade.

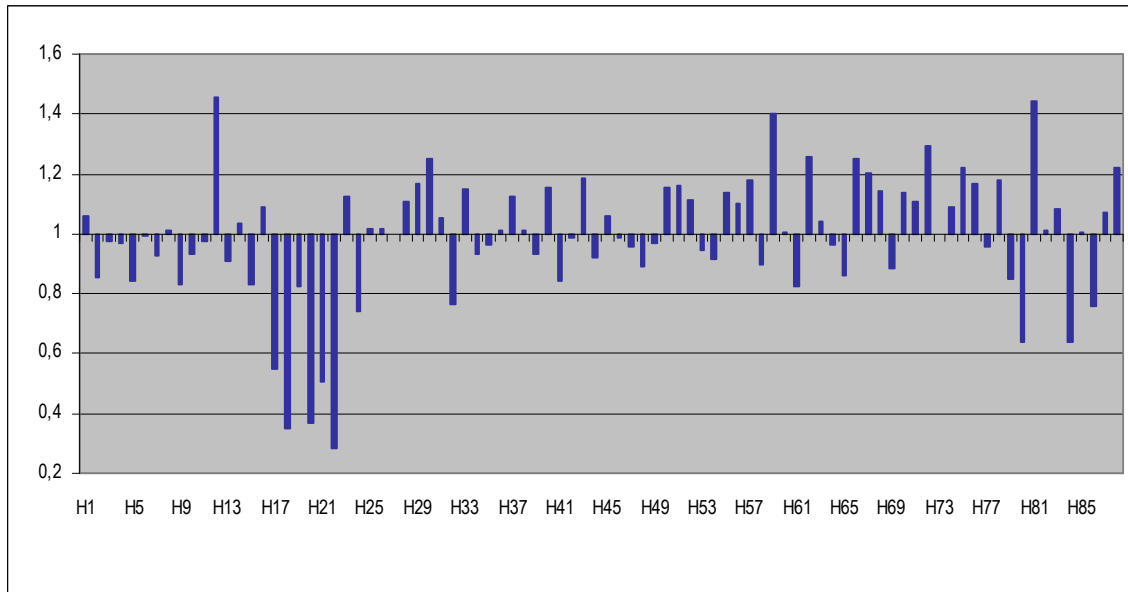
Neste indicador, um ratio superior a 1 significa que nesse hospital se observaram mais mortes que as esperadas. Assim, os hospitais nesta situação devem ser considerados como organizações de saúde com baixa efectividade. Pelo contrário um valor inferior a 1 traduz uma maior efectividade desse hospital.

Atendendo igualmente à fórmula de cálculo, valores mais afastados de 1 traduzem piores ou melhores efectividades, se este valor for respectivamente maior ou menor que 1.

Este indicador apresenta como principal virtude, o facto de se visualizarem de uma forma fácil e imediata a percentagem de mortes evitadas por cada hospital, quando o ratio é inferior a 1 e ainda a percentagem de óbitos que poderiam ser evitados, quando este valor é superior a 1.

Na Figura 33 são apresentadas as Taxas de Mortalidade Padronizadas por hospital (no Anexo X são apresentadas as taxas de mortalidade padronizadas por hospital).

Figura 33
Razão entre Taxas de Mortalidade Observadas e Esperadas (TMP)
por Hospital



Para os 88 hospitais em análise 46 apresentam valores superiores 1 (52%), sendo esta percentagem de 31% nos hospitais do Tipo I, 0% nos hospitais de Tipo II, 60% nos hospitais de Tipo III, 65% nos de Tipo IV e 67% nos de Tipo V.

Por Região de Saúde observa-se que na Região III 39% dos hospitais apresentam taxas de mortalidade superiores às esperadas, enquanto que estas percentagens são de 44%, de 63%, de 67% e de 100%, respectivamente para as Regiões V, IV, II e I.

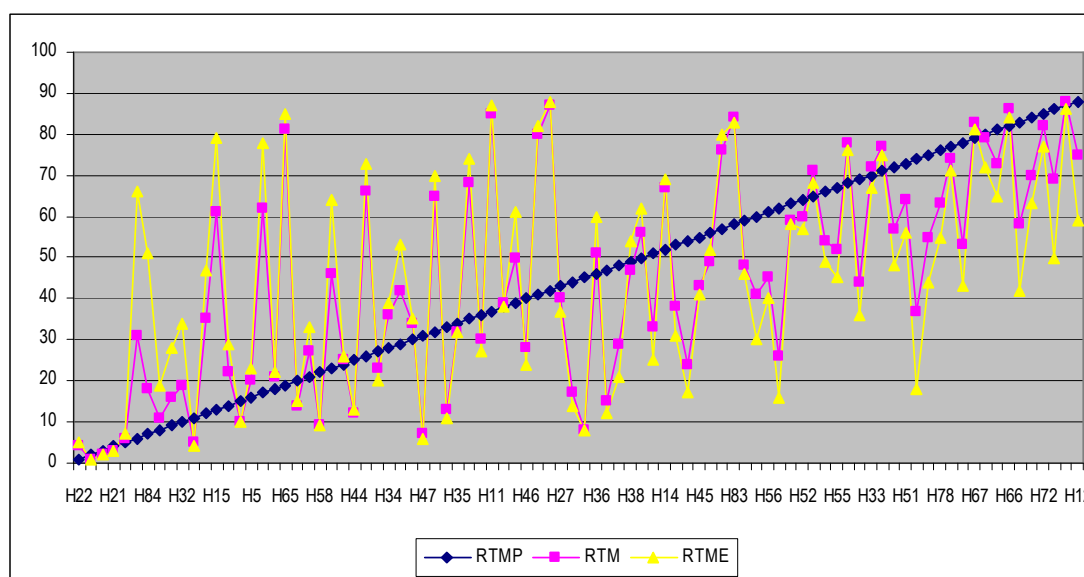
Na Figura 34 são apresentados as ordenações dos hospitais por ordem decrescente de efectividade. Nesta Figura são igualmente indicadas as ordenações dos hospitais quando se considera somente a taxa de mortalidade observada e a taxa de mortalidade esperada (ver Anexo X).

Podem fazer-se os seguintes comentários:

- Até à 42ª posição da ordenação da efectividade (ratios) os valores da taxa de mortalidade observada são iguais ou inferiores aos da taxa de mortalidade esperada;
- Nos 10 hospitais com maior efectividade 5 são do Tipo II, 2 dos Tipos III e IV e 1 do Tipo V. Por Região de Saúde, 4 são da Região III, 3 das Regiões IV e V;
- Na situação contrária, ou seja para os 10 hospitais com pior efectividade, encontram-se 5 do Tipo IV, e do Tipo III e 1 dos Tipos I e V. Destes

- hospitais, 5 estão localizados na Região de Saúde IV, 2 na Região I e 1 nas Regiões II, III e V;
- Não existe concordância no posicionamento dos hospitais em função do critério utilizado. Entre as ordenações da Taxa de Mortalidade Padronizada (TMP) e da Taxa de Mortalidade Observada o K de Cohen é de -0.011. Entre as ordenações da TMP e da Taxa de Mortalidade Esperada este coeficiente é de 0.000 e entre as ordenações das Taxas de Mortalidade Observada e Esperada esta estatística é de 0.069. A este propósito recorde-se que segundo Landis e Koch (1977) um valor de K inferior a 0.4 indica que a concordância é fraca;
 - Alguns exemplos ilustrativos na diferença na ordenação das organizações de saúde em função do critério utilizado podem ser encontrados nos hospitais 11, 15, 65, 73 e 79, para as comparações entre a Taxa de Mortalidade Padronizada e a Taxa de Mortalidade Observada, nos hospitais 15, 29, 65, 79 e 80 entre a Taxa de Mortalidade Padronizada e a Taxa de Mortalidade Esperada e nos hospitais 15, 29, 59, 80 e 84, para as Taxas de Mortalidade Observada e Esperada.

Figura 34
Ordenação dos Hospitais por Nível de Efectividade (RTMP)
Taxa de Mortalidade Observada (RTM) e Taxa de Mortalidade Esperada (RTME)

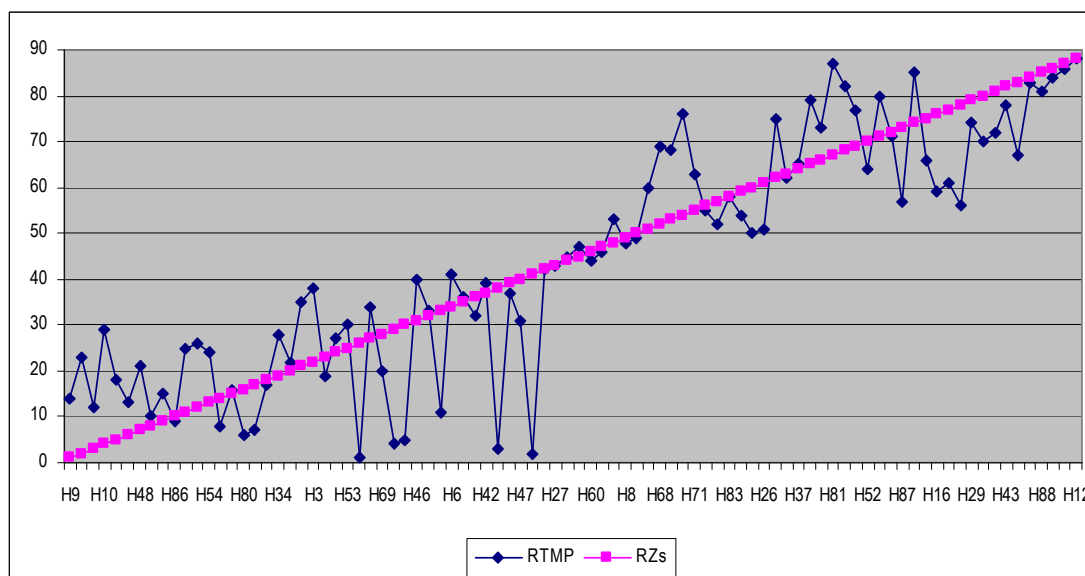


Contudo, como referem Shwartz, Ash e Iezzoni (1997) este método, a Taxa de Mortalidade Padronizada, embora constitua uma metodologia poderosa para se fazerem comparações entre hospitais, é, no entanto, um método indirecto que pode suscitar alguns problemas. Os problemas podem verificar-se quando as diferenças na casuística e na gravidade dos doentes entre instituições está presente, podendo constituir um factor de confundimento da análise.

Para tal sugerem uma alternativa de normalização/padronização directa, para se comparar mais correctamente o desempenho dos hospitais. Neste sentido, sugerem o cálculo de um “z score” de acordo com a equação apresentada na Metodologia.

Na Figura 35 são apresentados as ordenações dos hospitais por ordem decrescente de efectividade, sendo esta medida pelo “z score”. Nesta Figura são igualmente indicadas as ordenações dos hospitais quando se considera a Taxa de Mortalidade Padronizada (para se consultarem os valores ver Anexo X).

Figura 35
Ordenação dos Hospitais por Nível de Efectividade
Padronização Directa (“z score”) (RZs) e Indirecta (RTMP)



Em relação à Figura 35 evidenciam-se os seguintes aspectos:

- A identificação e ordenação dos hospitais com maior efectividade pelo método directo (“z score”) é completamente diferente da disponibilizada quando se utiliza a Taxa de Mortalidade Padronizada (TMP). O hospital com maior efectividade, o número 9, estava classificado somente na 14ª posição quando se utilizava a TMP, enquanto que os hospitais 13, 61, 10 e 62 que ocupam as posições seguintes quando se utiliza o “z score”, ocupavam respectivamente as 23ª, 12ª, 29ª e 18ª posições na Taxa de Mortalidade Padronizada. Deve ainda referir-se que nos dez primeiros classificados por esta nova metodologia, somente os hospitais 86 e 32 aparecem também entre os dez primeiros na classificação baseada na TMP;
- Nos 10 hospitais com maior efectividade (método directo) 5 são do Tipo I, 4 do Tipo III e 1 do Tipo V, situação que mais uma vez ilustra o facto de a informação disponibilizada ser completamente diferente da evidenciada

pela Taxa de Mortalidade Padronizada. Por Região de Saúde, 6 hospitais são da Região III e 2 das Regiões IV e V;

- Nos 10 hospitais com pior efectividade (método directo), embora 5 sejam comuns com a ordenação disponibilizada pela Taxa de Mortalidade Padronizada, deve igualmente concluir-se que as duas formas de classificação produzem resultados diferentes. Com o método directo encontram-se 8 hospitais do Tipo III e 1 dos Tipos I e V. Destes hospitais, 6 estão localizados na Região de Saúde IV e 1 em cada uma das restantes Regiões de Saúde;
- Finalmente é de referir que todos estes elementos são globalmente validados pela estatística de k de Cohen, com um valor de 0.034, indicando não concordância, ou seja os métodos directo e indirecto conduzem a ordenações completamente diferentes dos hospitais.

A disparidade encontrada nos resultados em função da utilização dos métodos indirecto ou directo para se avaliar a efectividade dos hospitais, relança a importância da consideração da casuística e da severidade do estado do doente para a comparação entre valores observados e esperados, pelo que à semelhança de diversos estudos efectuados por Iezzoni e outros (1996e), por Daley (1997) e por Iezzoni (1997c) serão utilizados os métodos directos, ou seja o “z score” para se avaliar o desempenho dos hospitais.

Esta questão, utilização do “z score” e a consequente alteração do panorama da avaliação dos hospitais, conforme se utilizam métodos directos ou indirectos, lança para discussão outro assunto: o facto de se utilizarem métodos directos para avaliar o desempenho dos hospitais pode ou não evitar a necessidade de se realizarem análises por tipo e por dimensão dos hospitais.

Este aspecto é parcialmente validado para a análise da efectividade merecendo igual tratamento para a eficiência e, na eventualidade de se encontrarem resultados similares pode conduzir a novas perspectivas sobre a avaliação de hospitais em Portugal, reposicionando a avaliação tradicional, com predomínio dos critérios de oferta, para uma análise em sede de procura.

Ou seja, ao contrário do que é tradicionalmente feito em Portugal, em que se procura definir “clusters” de hospitais com afinidades em termos de oferta, para se definirem grupos de comparação, se passe a comparar todos os hospitais em função das características da sua produção, independentemente, da sua localização, da respectiva dimensão e do seu tipo.

Depois de se ter realizado a avaliação do desempenho dos hospitais em função do total de doentes tratados, interessa agora proceder a uma maior desagregação. Num primeiro momento será feita uma análise por doentes cirúrgicos e médicos, para de seguida se centrar a análise por doença.

Em termos de casos cirúrgicos e médicos será utilizado também o “z score” como forma de ordenar os hospitais, embora, somente para efeitos de comparação, também se facultem informações sobre a ordenação correspondente à utilização da Taxa de Mortalidade Padronizada. Para qualquer uma das perspectivas serão retirados da análise os hospitais com menos de 100 doentes tratados (os hospitais 11 e 73 para os casos cirúrgicos).

Nas Figuras 36 a 39 são apresentados (para se consultarem os valores ver Anexo X):

- As ordenações para o método indirecto (Taxa de Mortalidade Padronizada) para todos os doentes (RTMPTOT) e casos cirúrgicos (RTMPCIR) e médicos (RTMPMED) (Figura 36);
- As ordenações para o método directo (“z score”) para todos os doentes (RZsTOT) e casos cirúrgicos (RZsCIR) e médicos (RZsMED) (Figura 37);
- As ordenações pelos métodos directo (RZsCIR) e indirecto (RTMPCIR) para os casos cirúrgicos (Figura 38);
- As ordenações pelos métodos directo (RZsMED) e indirecto (RZsMED) para os casos médicos (Figura 39).

Figura 36
Ordenações pelo Método Indirecto
Total de Casos e Casos Cirúrgicos e Médicos

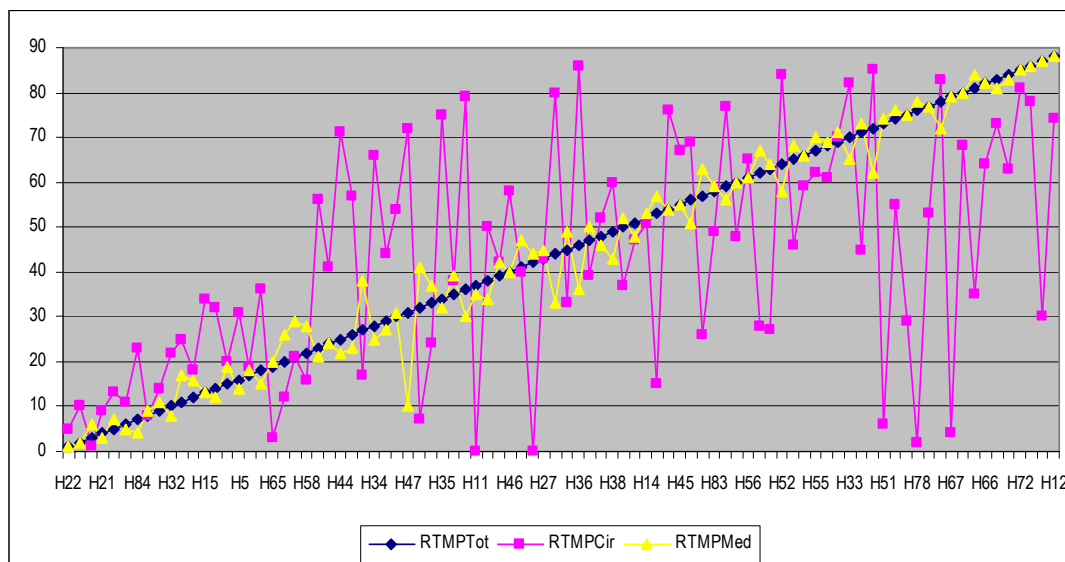


Figura 37
Ordenações pelo Método Directo
Total de Casos e Casos Cirúrgicos e Médicos

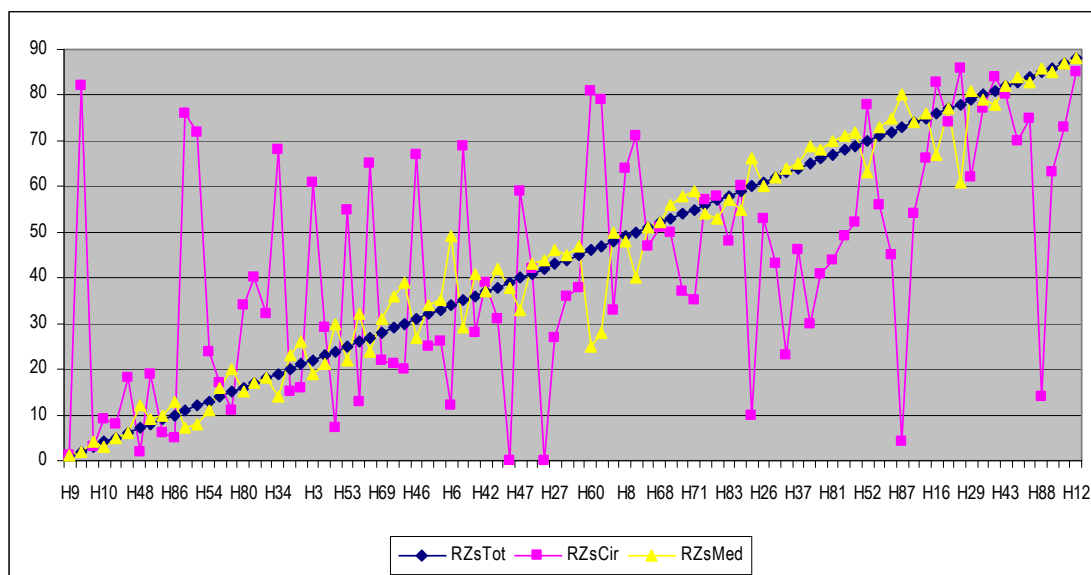


Figura 38
Ordenações pelos Métodos Directo e Indirecto
Casos Cirúrgicos

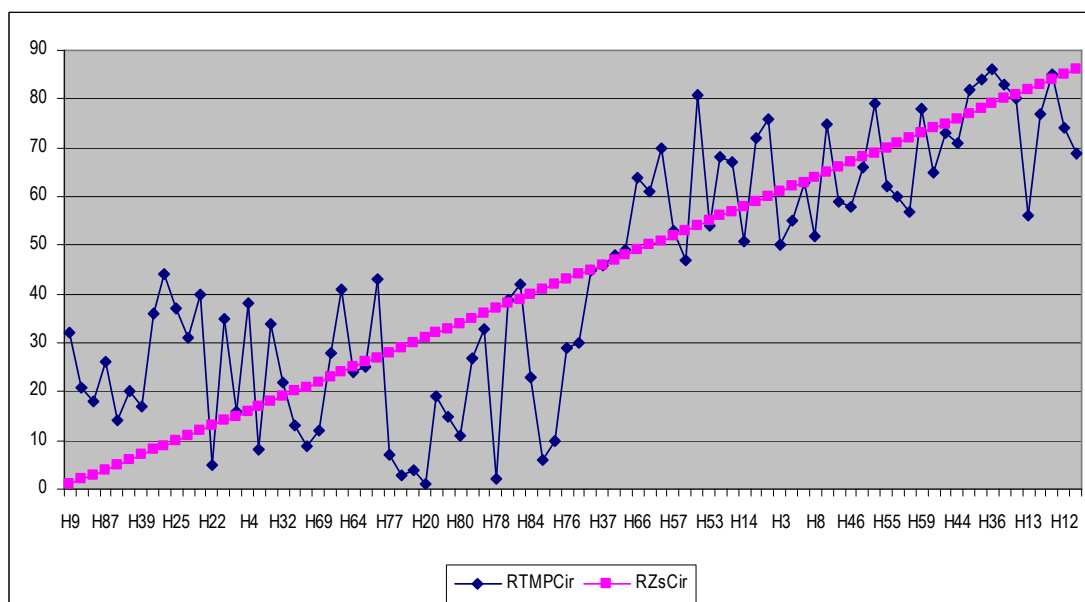
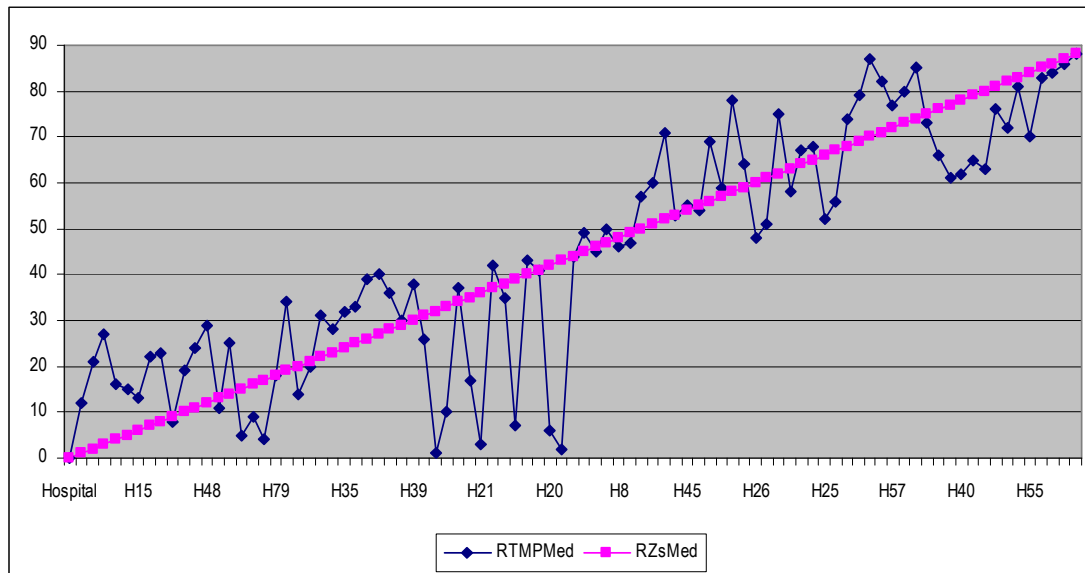


Figura 39
Ordenações pelos Métodos Directo e Indirecto
Casos Médicos



Utilizando a padronização indirecta (Taxa de Mortalidade Padronizada) observa-se que não existe concordância no desempenho dos hospitais entre o total de casos e os casos cirúrgicos (K de Cohen igual a 0.024), entre o total de casos e os casos médicos ($K=0.212$) e entre os casos cirúrgicos e os casos médicos ($K=0.024$).

Com a padronização directa ("z score") não existe igualmente concordância entre o total de casos e os casos cirúrgicos ($K=0.024$), entre o total de casos e os casos médicos ($K=0.165$) e entre os casos médicos e os casos cirúrgicos ($K=0.035$).

Estes valores permitem as seguintes observações:

- O desempenho dos hospitais é completamente diferente entre casos cirúrgicos e médicos e naturalmente quando estes são comparados com o total de doentes tratados;
- Embora o K de Cohen atinja sempre valores inferiores a 0.4, o que configura concordâncias fracas (Landis e Koch, 1977), deve referir-se que os seus valores são sempre superiores quando se comparam os hospitais entre o total de doentes e os casos médicos.

Para os casos cirúrgicos, à semelhança do que ocorreu para o total de doentes, não existe concordância na ordenação dos hospitais entre os métodos directo e indirecto, assumindo o K de Cohen um valor de 0.024.

Para os casos médicos, mais uma vez não existe concordância entre os métodos directo e indirecto para se avaliar a efectividade dos hospitais, sendo o K de Cohen de 0.035.

Pelas razões referidas, essencialmente pelo facto de o método directo tomar em consideração a gravidade dos doentes, será este o método que se irá utilizar para caracterizar o desempenho dos hospitais.

No Quadro XXXVIII são apresentados os 10 melhores e os 10 piores hospitais no que se refere à efectividade para os casos cirúrgicos. Paralelamente serão identificados o Tipo de hospital a que pertencem, bem como a respectiva localização em termos de Região de Saúde. A título ilustrativo serão indicadas as posições que ocupam para o total de doentes e para os casos médicos.

Quadro XXXVIII
Ordenação dos Hospitais por Efectividade (10 melhores e 10 piores)
Casos Cirúrgicos e Comparação com o Total de Casos e os Casos Médicos

Hospitais	Tipo	Região	Posição Casos Cirúrgicos	Posição Total de Casos	Posição Casos Médicos
9	I	III	1	1	1
48	III	V	2	7	12
61	III	III	3	3	4
87	V	IV	4	73	80
86	V	III	5	10	13
41	III	III	6	9	10
39	III	V	7	24	30
2	I	III	8	5	5
10	I	V	9	4	3
25	III	IV	10	60	66
33	III	IV	77	80	79
52	III	I	78	70	63
36	III	III	79	47	28
43	III	III	80	82	82
60	III	V	81	46	25
13	I	IV	82	2	2
16	I	IV	83	76	67
40	III	II	84	81	78
12	I	IV	85	88	88
1	I	V	86	78	61

Nos 10 hospitais com maior efectividade (casos cirúrgicos) 5 são do Tipo III, 3 do Tipo I e 2 do Tipo V. Destes 5 estão situados na Região de Saúde III, 3 na Região V e 2 na Região IV.

Para os 10 melhores hospitais (casos cirúrgicos), 7 surgem igualmente na ordenação para o total de doentes tratados, com a particularidade de o hospital 9 ocupar sempre a 1ª posição. Merece igualmente referência a situação do hospital 87, o 4º melhor para os casos cirúrgicos e somente o 73º para o total de doentes e ainda, embora com menor expressão, o observado nos hospitais números 25 e 39, com desempenhos distintos em função da perspectiva de avaliação.

A comparação entre os casos cirúrgicos e médicos produz igualmente ordenações distintas, com a existência de somente 5 hospitais a aparecerem nas duas listas. Por outro lado, as situações dos hospitais 87, 25 e 39 são igualmente interessantes, visto que apresentam ordenações completamente distintas para os casos cirúrgicos e médicos. Neste particular merece destaque o hospital 87, visto que aparece entre os 10 melhores nos casos cirúrgicos e entre os 10 piores nos casos médicos.

Para os 10 hospitais com pior efectividade 6 são do Tipo III e 4 do Tipo I. Destes 4 estão localizados na Região de Saúde IV, 2 nas Regiões III e V e 1 nas Regiões I e II.

Nos casos cirúrgicos a comparação com o total de admissões permite afirmar que existem 4 hospitais comuns no mau desempenho e que o hospital 13 apresenta um desempenho completamente diferente visto que ocupa uma das 10 últimas posições nos casos cirúrgicos, sendo o 2º melhor para todas as admissões. É ainda de evidenciar a situação do hospital 12, visto que apresenta o pior desempenho para todas as admissões, panorama que se mantém para os casos médicos.

A comparação entre os casos cirúrgicos e médicos apresenta igualmente ordenações diferentes, visto que somente 3 hospitais são comuns no respectivo desempenho. Os hospitais 13, 60 e 36 são aqueles que ilustram melhor o desempenho distinto entre hospitais para casos cirúrgicos e para casos médicos.

Analizando agora o desempenho dos hospitais nos casos médicos, no Quadro XXXIX são apresentados os 10 hospitais com melhor e pior efectividade. Para facilitar as comparações são igualmente identificadas as suas posições para todos os doentes e para os casos cirúrgicos.

Para os casos médicos nos 10 hospitais com melhor efectividade 6 são do Tipo I e 4 do Tipo III, pertencendo 5 à Região de Saúde III, 3 à Região IV e 2 à Região V.

A comparação com o total de admissões permite afirmar que para os 10 hospitais com melhor desempenho nos casos médicos, 8 apresentam igual comportamento para o total de doentes tratados.

Quando se comparam casos médicos e cirúrgicos o panorama é diferente visto que somente 5 hospitais são comuns às duas perspectivas, observando-se inclusivamente um comportamento relativamente distinto nos hospitais 7, 13 e 44, nos quais o desempenho nos casos cirúrgicos é pior do que o apurado para os casos médicos.

Quadro XXXIX
Ordenação dos Hospitais por Efectividade (10 melhores e 10 piores)
Casos Médicos e Comparação com o Total de Casos e os Casos Cirúrgicos

Hospitais	Tipo	Região	Posição Casos Médicos	Posição Total de Casos	Posição Casos Cirúrgicos
9	I	III	1	1	1
13	I	IV	2	2	82
10	I	V	3	4	9
61	III	III	4	3	3
2	I	III	5	5	8
15	I	IV	6	6	18
44	III	V	7	11	76
7	I	IV	8	12	72
32	III	III	9	8	19
41	III	III	10	9	6
33	III	IV	79	80	77
87	V	IV	80	73	4
29	III	IV	81	79	62
43	III	III	82	82	80
30	III	I	83	84	75
55	III	IV	84	83	70
62	III	IV	85	86	63
88	V	V	86	85	14
59	III	IV	87	87	73
12	I	IV	88	88	85

Para os 10 hospitais com pior efectividade, 7 são do Tipo III, 2 do Tipo V e 1 do Tipo I, pertencendo 7 à Região de Saúde IV e 1 às Regiões I, III e V.

A comparação entre hospitais para os casos médicos e para o total de admissões apresenta comportamentos semelhantes, visto que 9 organizações são comuns nas duas perspectivas. No entanto, o hospital 87 ocupa a 80ª posição para os casos médicos e a 73ª posição para o total de doentes tratados.

Quando se compara o desempenho dos hospitais para os casos médicos e cirúrgicos observa-se que somente 3 organizações são comuns nas duas perspectivas e que os hospitais 87 e 88 têm níveis de efectividade completamente distintos.

Estes resultados merecem ainda algumas reflexões suplementares:

- Um hospital, o número 9, aparece sempre com o melhor desempenho no método de avaliação directo;
- Por sua vez o hospital 12 apresenta quase sempre o pior desempenho com a padronização directa. Nos casos cirúrgicos é o segundo pior;
- Os hospitais do Tipo III apresentam um comportamento muito heterogéneo, para qualquer das perspectivas de análise apresentadas, Total de Admissões, Casos Cirúrgicos e Casos Médicos, o que para além de mais uma vez evidenciar que a procura de características de oferta para comparar hospitais pode não ser o melhor caminho, suscita que no futuro este Tipo de hospitais deva merecer uma atenção especial;
- O mesmo tipo de comentários pode ser feito para a Região de Saúde IV, onde é frequente encontrar hospitais com altos e baixos níveis de efectividade, o que, no mínimo, sugere a necessidade de um reforço dos mecanismos de administração, regulação e avaliação;
- Existe uma grande heterogeneidade no comportamento dos hospitais entre casos cirúrgicos e médicos, com maior expressão nos hospitais 3, 7, 13, 25, 34, 36, 44, 60, 87 e 88 (destes 10 hospitais, 5 são do Tipo III). No entanto, também se deve afirmar que existem 14 hospitais (2, 9, 12, 18, 24, 33, 41, 42, 43, 45, 56, 61, 68 e 74) com ordenações praticamente idênticas para os casos cirúrgicos e para os casos médicos (destes 8 são do Tipo III).

A grande disparidade no desempenho dos hospitais no que se refere à efectividade, entre casos médicos e cirúrgicos, justifica uma maior desagregação no plano de análise deste atributo.

Assim, a efectividade dos hospitais irá ser avaliada em termos de diagnósticos principais. Para tal, as doenças serão agregadas pelo Grandes Agrupamentos de Doenças do Disease Staging.

Para qualquer das análises serão eliminados os hospitais que tenham tratado menos de 100 doentes e serão somente identificados os cinco hospitais com melhor e pior desempenho em termos de efectividade (ver Quadro XL).

Existem 35 hospitais que não aparecem nenhuma vez posicionados entre aqueles que apresentam o melhor ou o pior desempenho em termos de efectividade.

Para os hospitais com melhor efectividade observa-se que 34 organizações diferentes aparecem, pelo menos uma vez, entre aqueles que ocupam as cinco primeiras posições. Quando se consideram somente os hospitais que ocupam a primeira posição na efectividade este valor desce para 10.

No que respeita aos hospitais com pior efectividade existem 40 unidades diferentes para as cinco últimas posições e 14, somente para o último lugar.

Para os 53 hospitais em análise, deve ainda referir-se que, quando se consideram as cinco posições para se qualificar este fenómeno, existem 13 hospitais que somente são classificados com bom desempenho e que se observam 19 hospitais somente classificados com mau desempenho. Os restantes 21 hospitais são classificados simultaneamente com bom e mau desempenho em função da respectiva doença.

Quadro XL
Hospitais com Maior e Menor Efectividade
por Grandes Agrupamentos de Doenças

Grandes Agrupamentos de Doenças	Hospitais com maior Efectividade	Hospitais com menor Efectividade
Doenças da mama	13, 86, 2, 9 e 3	88, 87, 62, 56 e 59
Doenças do sistema nervoso central	10, 13, 44, 61 e 84	12, 1, 43, 30 e 16
Doenças cardiovasculares e do coração	9, 10, 2, 60 e 38	12, 59, 25, 33 e 34
Doenças dermatológicas	4, 86, 61, 65, e 6	36, 13, 12, 83 e 40
Doenças do ouvido, nariz e garganta	9, 86, 13, 2, e 48	88, 87, 51, 25 e 16
Doenças gastrointestinais	9, 13, 10, 15 e 61	55, 12, 59, 62 e 44
Doenças ginecológicas	2, 9, 86, 35 e 43	88, 87, 29, 26 e 37
Doenças hepatobiliares	9, 13, 61, 48 e 1	40, 55, 59, 25 e 12
Doenças hematológicas	4, 9, 26, 15 e 25	88, 59, 87, 23 e 30
Doenças Imunológicas	9, 15, 2, 7 e 4	26, 10, 25, 48 e 62
Doenças linfáticas	9, 10, 35, 1, e 6	88, 87, 55, 86 e 4
Doenças dos órgãos genitais masculinos	3, 34, 13, 48 e 6	87, 88, 26, 16 e 59
Doenças musculoesqueléticas	61, 22, 8, 48 e 7	1, 12, 13, 16 e 36
Doenças endócrinas e metabólicas	9, 13, 10, 15 e 44	62, 59, 30, 33 e 12
Doenças obstétricas	20, 8, 41, 46 e 40	10, 61, 26, 39 e 1
Doenças oftalmológicas	10, 13, 2, 34 e 1	7, 40, 3, 30 e 60
Doenças do sistema nervoso periférico	9, 2, 7, 3 e 1	26, 46, 54, 4 e 48
Doenças psiquiátricas	13, 15, 1, 9 e 26	62, 59, 43, 28 e 29
Doenças renais	5, 34, 25, 2 e 61	33, 88, 59, 62 e 55
Doenças respiratórias	61, 1, 2, 25 e 46	62, 55, 23, 26 e 59
Doenças multisistémicas	13, 9, 2, 61 e 25	88, 40, 12, 33 e 55
Doenças vasculares (excepto coração)	9, 5, 10, 88 e 13	4, 12, 1, 59 e 60
Doenças não especificadas	9, 87, 58, 10 e 46	26, 66, 79, 67 e 59
Outras doenças	87, 13, 61, 86 e 15	38, 1, 40, 3 e 10

Analisando com maior detalhe os hospitais com melhor desempenho, por doença, podem fazer-se os seguintes comentários:

- Para o melhor desempenho, entre os 10 hospitais referenciados, somente 5 (hospitais 4, 9, 10, 13 e 61) recebem mais que uma citação. Como principal curiosidade deve referir-se que, com excepção do hospital 4, são estas organizações de saúde que ocupam as 4 primeiras posições na ordenação global;
- Para os 34 hospitais que ocupam as cinco primeiras posições, 12 aparecem referenciados uma vez, 6 duas e três vezes e 2 quatro vezes. Somente 6 recebem cinco ou mais citações, sendo os hospitais 9, 13 e 2, aqueles que apresentam uma maior frequência, respectivamente de 15, 12 e 10;
- Estes aspectos parecem mais uma vez evidenciar que a desagregação da análise por doença dá indicações diferentes das apreciações mais globais, para o total de doentes ou para casos cirúrgicos e médicos, pelo que se sugere a necessidade de se apresentarem estes dois planos de análise para se avaliar o desempenho dos hospitais.

Para os hospitais com piores níveis de efectividade, por doença, deve referir-se o seguinte:

- Para os 14 hospitais que aparecem referenciados como aqueles que apresentam o pior desempenho por agrupamento de doenças, somente 4 recebem mais que uma citação. Dentro destes destacam-se os hospitais 88, 26 e 62, os quais apresentam a pior efectividade em 6, 3 e 3 Grandes Agrupamentos de Doenças, respectivamente. Apesar desta situação o hospital 88 ocupa a 85ª posição na ordenação global e o hospital 62 a 86ª posição da mesma ordenação, enquanto que o hospital 26 ocupa a 21ª posição;
- No que se refere às cinco últimas posições, 18 hospitais somente aparecem referenciados num grande agrupamento de doença, enquanto que 7 surgem em dois, 2 em três e 3 em quatro. Com cinco ou mais citações, existem 10 hospitais, devendo evidenciar-se os hospitais 59, 12 e 88. O primeiro, obtém 12 citações, enquanto que os restantes apresentam o pior desempenho em 9 e em 8 Grandes Agrupamentos de Doenças, respectivamente. Os hospitais 12 e 59 ocupam as duas últimas posições quando se utiliza a ordenação global de doentes.

A combinação de todas as perspectivas de análise apresentadas, merece ainda algumas reflexões suplementares, designadamente no que se refere à relativa dispersão dos hospitais na efectividade por doença, tanto para os melhores níveis, como para os piores e à existência de hospitais, tanto isoladamente, como por Tipo, que apresentam simultaneamente melhor e pior desempenho.

Em relação à dispersão dos hospitais com melhor desempenho, pode, por exemplo, referir-se o hospital 9, afinal o que apresenta o melhor desempenho, tanto na ordenação global, como nos casos médicos e cirúrgicos, o qual somente é o melhor em 10 dos 24 Grandes Agrupamentos de Doenças.

Para o pior desempenho deve, essencialmente, apontar-se o facto de somente 25 hospitais aparecerem entre as cinco últimas posições, uma ou duas vezes.

Deve ainda referir-se a situação de algumas organizações de saúde em que se observa a existência de bons e maus níveis de desempenho (21 hospitais), sendo as situações mais ilustrativas encontradas nos hospitais 1, 4 e 25. O hospital 1 é citado 6 vezes nas cinco primeiras posições e 5 vezes nas cinco últimas posições, enquanto que no hospital 4 tal é observado 3 vezes, tanto para as cinco primeiras, como para as cinco últimas posições e no hospital 25 a frequência é de quatro, igualmente para o bom e mau desempenho.

Quando se compara esta análise com a efectividade global dos hospitais, observa-se que não existe qualquer correspondência, dado que o hospital 1 ocupa a 78ª posição na ordenação global, o hospital 4 a 21ª posição e o hospital 25 a 60ª posição, o que parece mais uma vez comprovar a inexistência de uma concordância entre a efectividade medida em termos globais e em termos parciais.

Esta situação suscita, desde já, questões de dupla natureza.

Por um lado, parece reforçar o que anteriormente foi referido, tanto no que respeita à pesquisa de homogeneidade por critérios de oferta, visto que mesmo dentro do mesmo hospital existe uma grande variação no desempenho por doença, como na necessidade de se reforçarem os mecanismos de avaliação, de regulação e de administração do sector hospitalar português.

Por outro lado, a existência de mecanismos de liberdade de escolha por parte dos consumidores, pode suscitar exigências suplementares, visto que pode não ser suficiente a selecção de uma unidade de saúde, podendo os “consumidores” optar por especialidades/serviços de hospitais diferentes.

Este aspecto, pode por sua vez, criar dificuldades suplementares na política de financiamento hospitalar, bem como nos mecanismos de regulação e de avaliação, pelo que deverá merecer atenção crescente, tanto por parte dos investigadores, bem como dos responsáveis do sector de saúde.

5.2.2. Eficiência

A eficiência técnica traduz sempre uma relação entre recursos e produção ou entre produção e recursos (Pereira, 1993). Em termos globais, para os indicadores de eficiência podem ser considerados os custos médios e/ou a demora média (Butler, 1995).

Neste estudo e atendendo, por um lado, à inexistência de custos por produto e por doente nos hospitais portugueses e, por outro lado, à metodologia adoptada, utilização do ajustamento pelo risco, para se compararem valores observados com os esperados, torna-se impossível, pelo menos de uma forma directa, utilizar os custos como medida da eficiência técnica.

Neste sentido, a medida de eficiência que vai ser utilizada respeita à demora média. A demora média expressa uma relação entre a duração de internamento, a medida de recursos e os doentes saídos, a medida da produção.

Para além dos aspectos referidos anteriormente sobre as limitações da utilização desta medida, convém, neste momento, evidenciar a questão da intensidade da utilização de recursos. Na realidade as avaliações que utilizam a demora média como indicador de eficiência admitem, pelo menos implicitamente que, existe uma constância no nível de utilização de recursos, tanto entre doenças, como na severidade do estado do doente e ainda durante o próprio episódio de internamento (Costa, 1994 e Butler, 1995).

Como é do conhecimento geral, esta intensidade na utilização de recursos pode variar em função (Butler, 1995):

- Da própria doença, como por exemplo, no limite se podem citar as situações cirúrgicas e médicas;
- Da gravidade dos doentes, admitindo-se uma relação directa entre estas duas características;
- Da estadia do doente, com uma diminuição progressiva do nível de utilização de recursos com o aumento da duração de internamento.

Apesar destes problemas a utilização da demora média como indicador de eficiência técnica, não deve ser descartada, visto que para além de constituir uma medida de rotina dos hospitais, pode ainda ser considerada como relativamente fiável, visto que no limite os procedimentos para o seu apuramento são idênticos entre os diversos hospitais (Iezzoni, et al 1996c e Schwartz, et al, 1996).

Por outro lado, contém ainda algumas indicações úteis e pertinentes sobre o desempenho dos hospitais, essencialmente atendendo a que para os doentes uma diminuição da estadia, para o mesmo nível de qualidade dos cuidados prestados é uma expectativa e constitui ainda um resultado favorável e que para

os hospitais, incluindo os gestores e os prestadores, a diminuição das estadias, pode constituir uma forma de diminuir o consumo desnecessário de recursos e desta forma aumentar a eficiência (Philips et al, 2000).

Para se caracterizar e avaliar o desempenho dos hospitais e, à semelhança do que foi feito para a efectividade, vão ser utilizados três anos em detrimento de um, para que desta forma se atenuem eventuais efeitos de fenómenos conjunturais ou episódicos.

Mais uma vez, conforme se procedeu na avaliação da efectividade, tendo igualmente presente que eventuais diferenças na demora média podem traduzir diferenças nas características dos doentes será utilizado um sistema de ajustamento pelo risco – o Disease Staging – para desta forma se identificarem e medirem as principais diferenças na duração do internamento dos hospitais, em função dos valores observados e dos previstos pelo referido sistema de classificação de doentes.

Assim, o presente capítulo é constituído por duas partes:

- Identificação e comparação da demora média nos hospitais;
- Avaliação da eficiência dos hospitais.

5.2.2.1. Demora Média

A duração do internamento hospitalar é conhecida por ter uma assimetria positiva (Kulinskaya, Staudte e Hales, 2002; Clark e Ryan, 2002; Shwartz et al, 1996 e lezzoni et al, 1996c). Tal pode ser devido a um diferente comportamento entre os doentes com alta vivos e mortos, à própria doença, à severidade do estado do doente a ainda a diferentes padrões de tratamento por parte dos hospitais (Shwartz et al, 1996 e lezzoni et al, 1996c).

Tal suscita que por vezes se identifiquem “outliers”, ou seja casos com duração de internamento anormalmente longa ou curta ou que, ainda se procedam a transformações na duração de internamento, essencialmente através da utilização do logaritmo natural (Ln) da mesma (Shwartz et al, 1996 e Clark e Ryan, 2002).

Com este procedimento pretende-se que a distribuição da duração de internamento tenha uma maior aproximação da distribuição normal e desta forma tornar mais robustas as tradicionais análises com estatísticas paramétricas.

Para este estudo para além de uma primeira recodificação de todos os casos com duração de internamento igual a zero no resumo informatizado de alta, para uma duração de internamento igual a 1, torna-se igualmente necessário definir o melhor procedimento para a definição da população em estudo para efeitos de avaliação da eficiência.

Para os 2315189 doentes em estudo a distribuição da duração de internamento apresenta uma forte assimetria positiva (“skewness” = 14.8 e erro padrão = 0.002). A utilização do logaritmo natural da duração de internamento, embora atenua substancialmente este enviesamento (“skewness” = 0.33 e erro padrão = 0.002), não evita que o fenómeno esteja presente.

Daqui, ser ainda necessário identificar a existência de eventuais “outliers”, tendo em vista a sua eliminação da população em estudo. Para tal podiam-se utilizar os limiares inferiores e superiores definidos na Portaria nº189/2001 de 2 de Março, ou utilizar técnicas mais robustas para o efeito, essencialmente através da amplitude entre quartis (Emerson e Strenio, 1983), tendo-se optado por esta última alternativa (ver Metodologia).

Ficam assim em análise 2176216 doentes, com a particularidade de se observar uma grande diminuição na assimetria da distribuição da duração de internamento (“skewness” = 3.37) e para o logaritmo natural da duração de internamento (“skewness” = 0.208).

Para os três anos em análise a demora média foi de 6.10 dias, com a particularidade de se manter praticamente constante ao longo de período, sendo de 6.09, 6.11 e 6.10 dias, respectivamente para 1999, 2000 e 2001.

No Quadro XLI são apresentadas as demoras médias globais, por Região de Saúde e por Tipo de hospital.

Quadro XLI
Demoras Médias por
Região de Saúde e Tipo de Hospital

Região de Saúde / Tipo de Hospital	Demora Média Região de Saúde	Demora Média Tipo de Hospital
I	6.57	6.63
II	6.35	4.13
III	6.14	5.83
IV	6.27	6.02
V	5.81	7.23
Continente	6.10	6.10

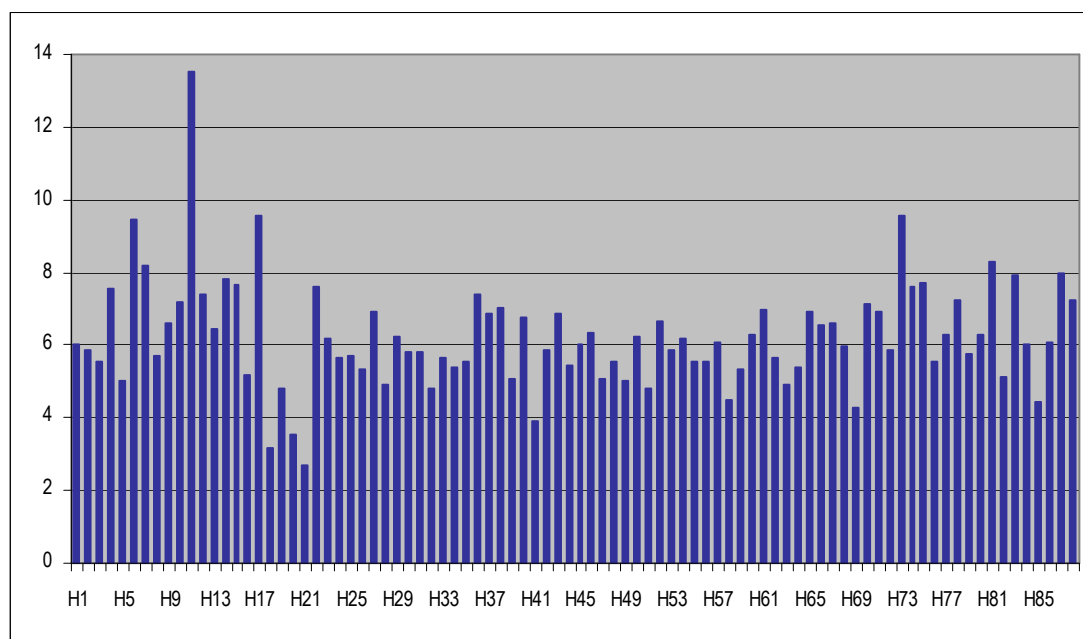
Fonte: Direcção Geral da Saúde

As Regiões de Saúde com demoras médias mais baixas são a V e a III, enquanto que a I e a II são as que apresentam valores mais elevados.

Por Tipo de hospital, observa-se que são os do Tipo II e III que apresentam demoras médias mais baixas, enquanto que na situação contrária se encontram nos hospitais do Tipo V e I.

Na Figura 40 é representada a demora média por hospital (no Anexo XI apresentam-se os valores da demora média por hospital).

Figura 40
Demora Média por Hospital



A demora média por hospital oscila entre os 2.71 dias no hospital 21, até os 13.55 dias no hospital 11. Contudo, existe uma grande homogeneidade entre a demora média dos hospitais, visto que o coeficiente de variação é de 0.24.

Outro aspecto a salientar respeita à forte associação entre demoras médias e Taxas de Mortalidade por hospital, o coeficiente de correlação é de 0.617 (significativo a 0.01), o que parece evidenciar que a demora média e a taxa de mortalidade apresentam o mesmo sentido, ou seja que, quanto maior a taxa de mortalidade observada maior a duração de internamento.

Este aspecto aqui indiciado pode apresentar alguma importância, pelo facto de a duração de internamento e a mortalidade apresentarem um comportamento semelhante, o que parece contrariar algumas conclusões de estudos internacionais (Shwartz et al, 1986 e Clark e Ryan, 2002), por exemplo onde é admitido que a relação entre estes dois aspectos pode ser quadrática, com durações de internamento mais baixas, tanto para níveis de severidade mais baixos, como para os mais elevados.

Parecendo ainda justificar as opções de alguns sistemas de classificação de doentes, como por exemplo, os All Patients Refined Diagnosis Related Groups ou os International Refined Diagnosis Related Groups, em que os escalões de

gravidade são estabelecidos em função da respectiva duração de internamento (Averill et al, 1998 e Mullin, Averill e Boucher, 2002).

Atendendo à importância do assunto, o mesmo será discutido num momento posterior deste estudo.

Em seguida serão apresentadas as estimativas para a duração de internamento, recalibradas aos dados portugueses e, desta forma, analisar a eficiência dos hospitais.

5.2.2.2. Avaliação da Eficiência

Conforme foi referido a avaliação da eficiência vai ser feita com recurso aos valores da demora média observados e aos estimados pelo “software” do Disease Staging, ponderados por um valor de referência (ver Metodologia).

Atendendo a que o Disease Staging inclui três escalões ordinais de gravidade, a comparação entre valores observados e estimados será feita, por escalão de gravidade, bem como para a totalidade dos doentes.

A principal razão para este procedimento pode ser encontrada no facto de o desempenho dos hospitais ser ou não idêntico, para a globalidade dos doentes tratados e para cada estadio de severidade.

Mais uma vez, à semelhança do realizado na avaliação da efectividade, os apuramentos serão feitos para todos os doentes, para os casos cirúrgicos e médicos e para agrupamentos de doenças.

Para esta apreciação da eficiência são candidatos todos os hospitais, com excepção daqueles que, para a respectiva perspectiva, não atinjam os 100 doentes tratados.

A avaliação da eficiência será feita pelo método indirecto, ou seja, a divisão entre valores observados e esperados para a demora média e pelo método directo – utilização de um “z score”, tendo em vista a padronização da duração de internamento de todos os hospitais.

Embora este “z score” possa ser calculado de diversas formas, a fórmula escolhida é idêntica à utilizada para a efectividade, para facilitar a compreensão da análise da eficiência hospitalar, bem como para se tornarem mais directas as comparações com a efectividade dos cuidados prestados por hospital (ver na Metodologia a forma de cálculo desta estatística).

No Quadro XLII são apresentadas as Demoras Médias Observadas (DMOT) e Esperadas (DMET) por Região de Saúde e por Tipo de hospital, bem como a razão entre as respectivas demoras médias e o “z score”.

Quadro XLII
Demoras Médias Observadas e Esperadas
por Região de Saúde e por Tipo de Hospital, Razão e “Z score”

	Região de Saúde				Tipo de Hospital			
	DMOT	DMET	Razão	Z score	DMOT	DMET	Razão	Z score
I	6.57	6.19	1.061	0.016	6.63	6.66	0.996	-0.007
II	6.35	5.89	1.078	0.013	4.13	4.08	1.011	0.001
III	6.14	6.06	1.014	0.024	5.83	5.86	0.996	-0.008
IV	6.27	6.38	0.983	-0.040	6.02	5.88	1.025	0.008
V	5.81	5.85	0.993	-0.013	7.23	6.97	1.037	0.006
Total	6.10	6.10			6.10	6.10		

Para as Regiões IV e I são estimadas as demoras médias mais elevadas e para as Regiões V e II os valores mais baixos. Estes resultados são diferentes das demoras médias observadas, visto que, por exemplo, a Região II é a que apresenta o segundo valor mais elevado. A Região IV é a mais eficiente (segundo o método directo) e as Regiões III e I as menos eficientes.

Por Tipo de hospital, são os do Tipo V e I que apresentam estimativas mais elevadas para a demora média e os hospitais do Tipo II e III apresentam os valores mais baixos. Por sua vez os hospitais dos Tipos III e I são os mais eficientes, enquanto que os dos Tipos IV e V são os menos eficientes.

Figura 41
Demoras Médias Observadas e Esperadas por Hospital
Estadio 1

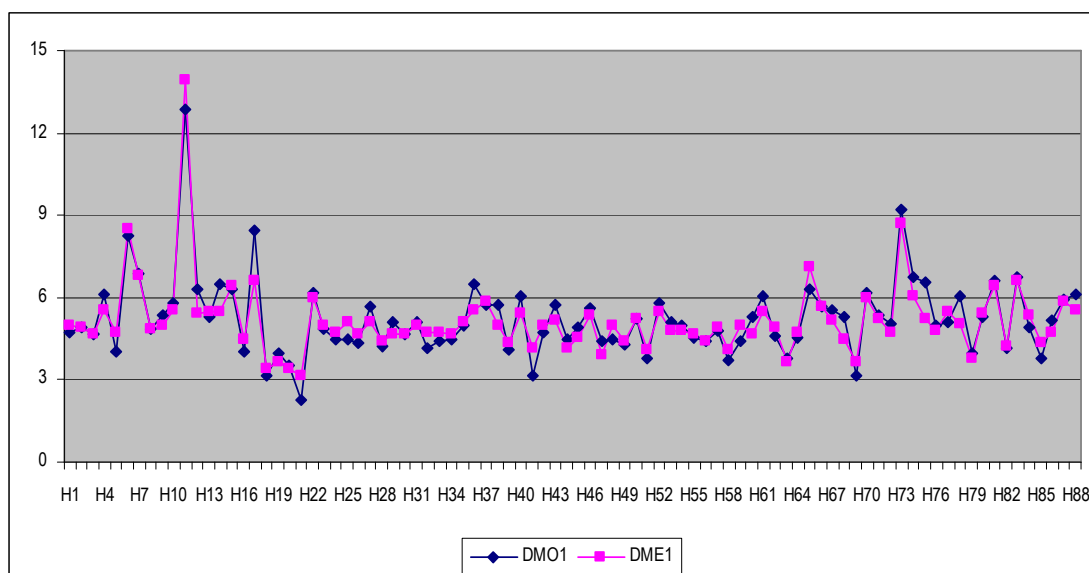


Figura 42
Demoras Médias Observadas e Esperadas por Hospital
Estadio 2

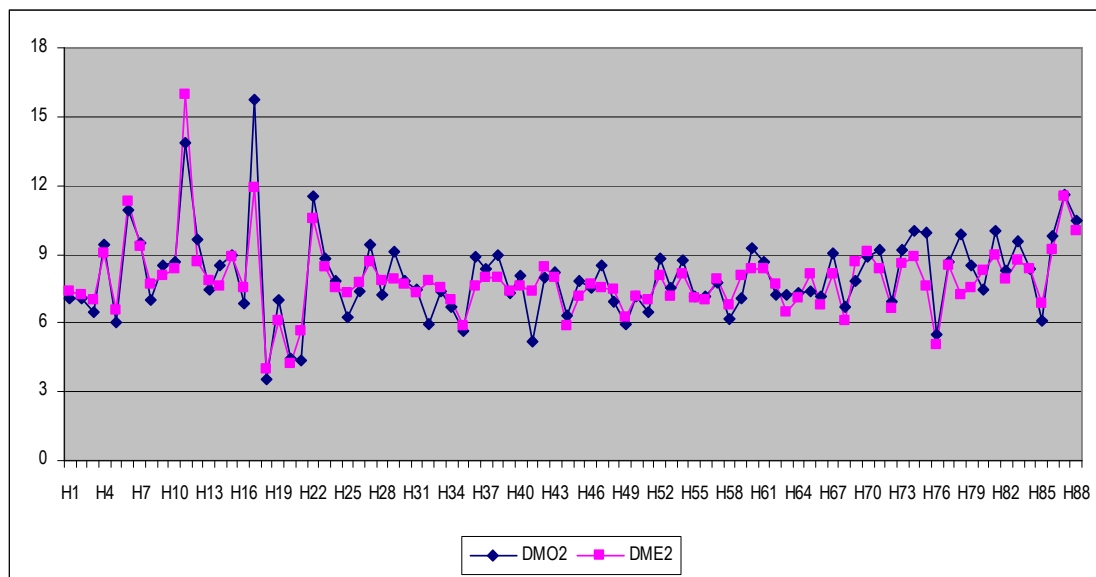


Figura 43
Demoras Médias Observadas e Esperadas por Hospital
Estadio 3

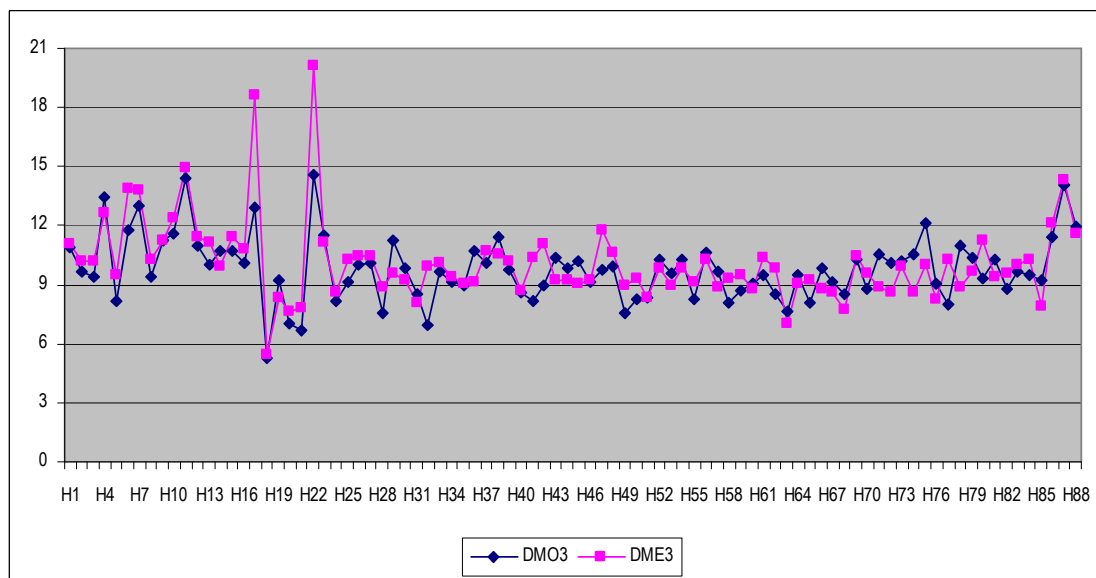
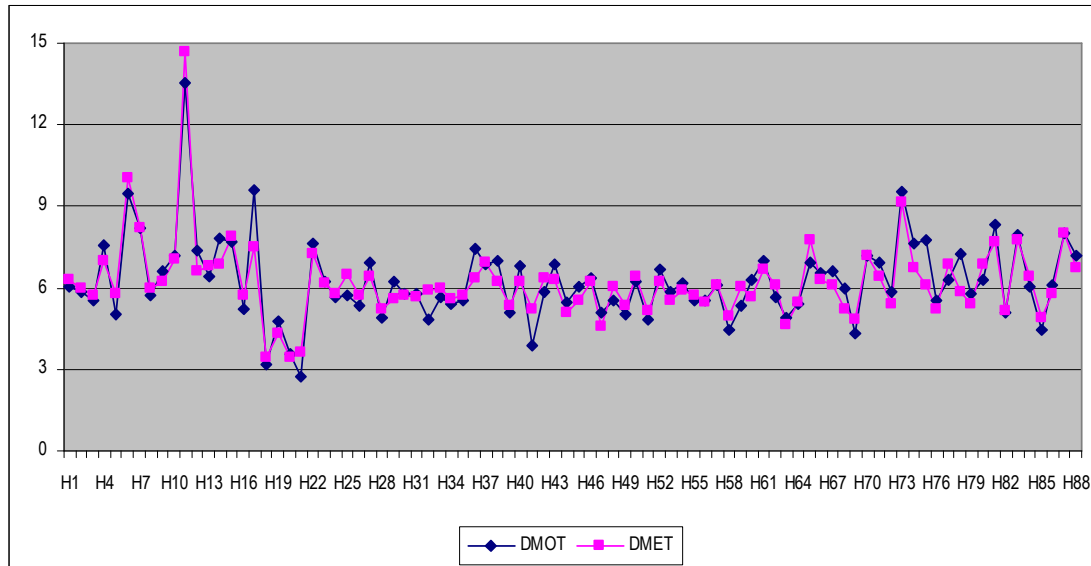


Figura 44
Demoras Médias Observadas e Esperadas por Hospital
Total de Doentes



Nas Figuras 41 a 44 foram apresentadas as demoras médias observadas esperadas por hospital e por estadio de severidade e para a totalidade dos doentes. No Anexo XI podem ser consultados os respectivos valores.

Em relação à demora média esperada para todos os doentes observa-se uma dispersão pequena, com o coeficiente de variação de 0.23, com um valor máximo no hospital 11 (14,66 dias) e o mínimo no hospital 18 (3.42 dias). É ainda de referir que a heterogeneidade entre hospitais é praticamente igual à da demora média observada, cujo coeficiente de variação é de 0.24.

Analisando por estadios de gravidade pode afirmar-se o seguinte:

- Para o estadio 1, a demora média esperada mais elevada é encontrada no hospital 11 (13.95 dias) e a mínima no hospital 21 (3.13 dias). Deve ainda referir-se que não existem diferenças significativas entre demoras médias esperadas e observadas por hospital dado que os coeficientes de variação são, respectivamente de 0.26 e de 0.27;
- Para o estadio 2, o máximo é mais uma vez encontrado no hospital 11 (16.00 dias) e o mínimo no hospital 18 (3.96 dias). Mais uma vez o coeficiente de variação é menor nos valores esperados (0.20), do que na demora média observada (0.23);
- Para o estadio 3, o máximo é encontrado no hospital 22 (20.08 dias) e o mínimo no hospital 18 (5.48 dias), embora com maior homogeneidade nos valores observados da demora média, sendo o coeficiente de variação de 0.17, enquanto que para a demora média esperada esta estatística é de 0.21.

A discussão sobre se a utilização de estadios de severidade introduz ou não informações distintas das disponibilizadas pelo total de doentes pode ser feita através da comparação das ordenações dos hospitais para cada um dos estadios de gravidade e para o global dos doentes, tanto para a demora média observada, como para a esperada.

Esta análise irá ser feita através da estatística *K de Cohen*, cujos valores são apresentados no Quadro XLIII.

Quadro XLIII
Concordância na Atribuição de Ordenações por Hospital (*K de Cohen*)
para as Demoras Médias Observadas e Esperadas por Estadio
e para o Total dos Doentes

	DMO2	DMO3	DMOT	DME1	DME2	DME3	DMET
DMO1	0.011	0.011	0.069	0.092			
DMO2		0.034	0.034		0.057		
DMO3			0.011			0.011	
DMOT							0.057
DME1					0.000	0.000	0.115
DME2						0.034	0.034
DME3							0.011

DMO: Demora Média Observada por Estadios 1, 2, 3 e Total;

DME: Demora Média Esperada por Estadios 1, 2, 3 e Total

Para a demora média observada pode referir-se que os estadios de severidade estão a dar diferentes informações entre si e em relação ao total de doentes tratados por hospital. Na realidade, o *K de Cohen* é sempre inferior 0.4, o que segundo Landis e Koch (1977) configura concordâncias fracas. Mesmo assim é de referir que o valor mais elevado na ordenação dos hospitais por demora média é obtido entre o estadio de gravidade 1 e o total de doentes.

Para as demoras médias esperadas, o panorama é idêntico, ou seja concordância fraca. No entanto deve referir-se a comparação das ordenações dos doentes no estadio 1 com os restantes doentes, visto que o *K de Cohen* é zero para os estadios 2 e 3 e de 0.115 (o valor mais elevado) para a totalidade dos doentes.

Estes valores vêm confirmar o interesse da avaliação da eficiência, tanto em termos globais, como por estadio de gravidade.

A comparação entre valores observados e esperados, por sua vez, também contém informações distintas. De facto, tanto para os estadios de severidade, como para a globalidade dos doentes, a ordenação dos hospitais é completamente diferente, com todos os *Ks de Cohen* a apresentarem valores inferiores a 0.1, mesmo assim com um máximo para os doentes do estadio 1

(0.092) e um mínimo para os doentes no escalão de severidade 3 (0.011), embora estes valores tenham pouco significado.

Esta situação permite evidenciar dois aspectos: em primeiro lugar, a falta de precisão quando se compara o desempenho dos hospitais somente em função do que foi feito, neste particular a demora média; e em segundo lugar, a importância do ajustamento pelo risco para avaliar a eficiência hospitalar.

Este facto para além de entrar em consideração com as características dos doentes para avaliar a demora média, vem diminuir ou mesmo eliminar os argumentos apresentados pelos hospitais, essencialmente pelos seus gestores e profissionais, quando pretendem justificar eventuais resultados mais fracos, através das características específicas e próprias das suas organizações e dos seus doentes.

Ainda dentro deste plano de análise convém avaliar se, tanto os resultados observados, como os esperados, estão ou não associados com a gravidade dos doentes e ainda, na eventualidade desta associação existir, se são os doentes mais ou menos graves que contribuem para melhores desempenhos.

Utilizando a informação constante no Quadro XLIII, a análise que irá ser feita pretende relacionar o desempenho global dos hospitais, medido quer pela demora média observada, quer pela demora média esperada, com o desempenho por nível de severidade. Assim, irão ser feitas comparações da ordenação dos hospitais, pelas diferentes perspectivas, através do coeficiente de correlação de Spearman (ver Quadro XLIV).

Quadro XLIV
Associação entre as Ordenações Globais
e por Estádios de Severidade

	DMO1	DMO2	DMO3
DMOT	0.971	0.842	0.681

	DME1	DME2	DME3
DMET	0.956	0.759	0.580

DMO: Demora Média Observada por Estádios 1, 2, 3 e Total;

DME: Demora Média Esperada por Estádios 1, 2, 3 e Total

Tanto para os valores observados, como para os valores esperados, todas as correlações são significativas. É ainda de salientar que as correlações são sempre superiores nos valores observados.

Deve ainda referir-se que, pese embora o facto de todas as correlações serem significativas, o valor mais elevado surge sempre na associação entre a ordenação do total de doentes, com a dos doentes no estadio 1.

Este aspecto, merece algumas considerações suplementares, visto que os hospitais podem tentar melhorar o seu desempenho através de uma escolha de doentes que lhes são mais favoráveis, essencialmente os menos graves, o que para além dos eventuais impactes negativos sobre a população que é servida, pode ainda penalizar outras organizações, as quais numa perspectiva ampla podem ser consideradas como suas concorrentes.

Atendendo à importância deste aspecto na definição de prioridades em saúde, bem como nas metodologias de avaliação e de regulação, o mesmo será discutido posteriormente, nomeadamente através do cruzamento de informações e de avaliação do desempenho, tanto com a efectividade dos cuidados, bem como com a política de admissão dos hospitais.

Nas Figuras 45 e 46 são apresentadas as ordenações dos hospitais no seu desempenho na eficiência, utilizando os métodos indirecto (razão entre demora média observada e demora média esperada) e directo ("z score"), para os doentes dos estádios 1, 2 e 3 e para o total dos doentes tratados. No Anexo XI podem ser consultados os respectivos valores.

Figura 45
Ordenações dos Hospitais pelo Método Indirecto
Estádios 1, 2 e 3 e Total de Doentes

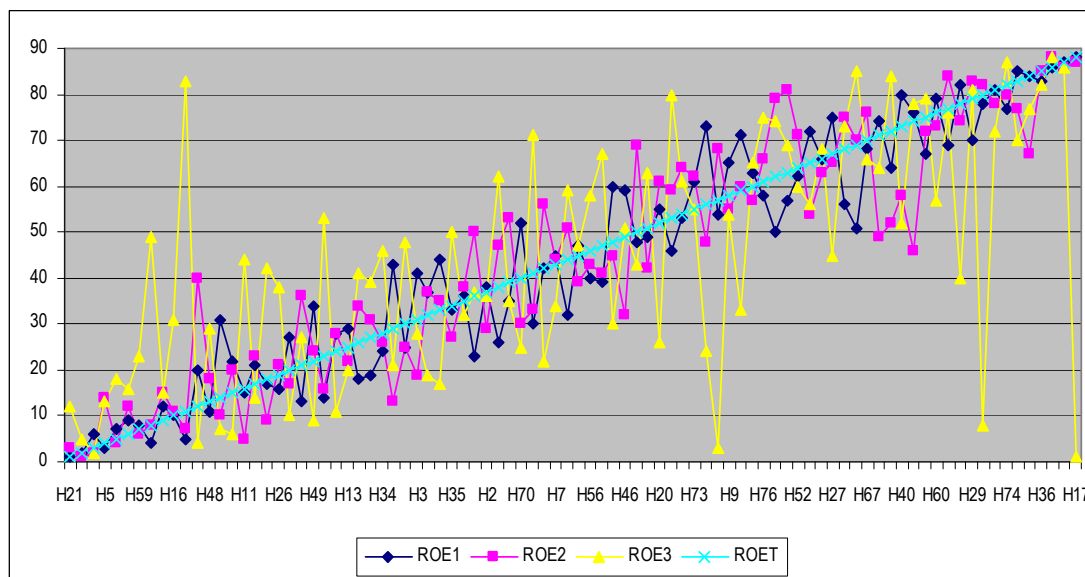
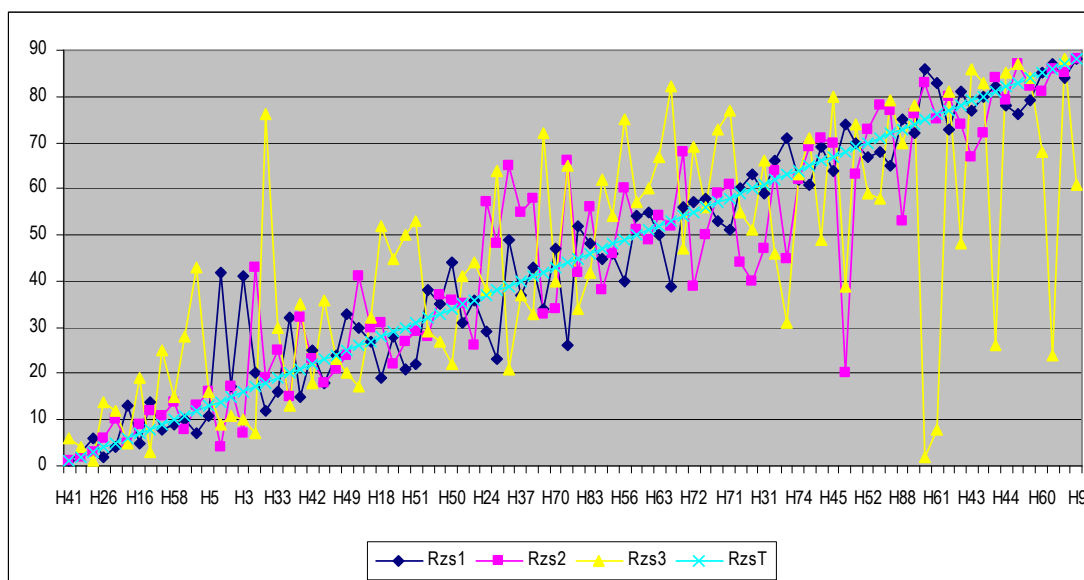


Figura 46
Ordenações dos Hospitais pelo Método Directo
Estádios 1, 2 e 3 e Total de Doentes



Mais uma vez a utilização dos métodos indirecto e directo conduzem a ordenações diferentes por parte dos hospitais. A título exemplificativo pode referir-se que para os 10 hospitais com melhor desempenho, medido pelo método directo ("z score"), somente 5 estão presentes quando avaliados pelo método indirecto, ou que para os 10 piores em termos de desempenho pelo método directo, somente 2 estão presentes pelo método indirecto.

Uma abordagem mais agregada, expressa através da concordância nas ordenações para todos os hospitais, o *K de Cohen* é de 0.023. Atendendo a que esta estatística é de 0.092, 0.046 e 0.023, respectivamente para os escalões de severidade 1, 2 e 3, pode concluir-se que os dois métodos disponibilizam informação completamente diferente.

Neste sentido, à semelhança do que foi realizado para a efectividade, opta-se pelo método directo para se ordenarem hospitais pela eficiência.

Nesta lógica, a utilização dos escalões de severidade disponibiliza indicações igualmente diferentes sobre o desempenho dos hospitais, sendo a concordância entre os escalões 1 e 2 de 0.057, entre os escalões 1 e 3 de 0.000 e entre os escalões 2 e 3 de 0.023. Em relação ao total da produção o *K de Cohen* é de 0.069, 0.057 e 0.011, respectivamente para os estádios 1, 2 e 3.

Por outro lado, nos 10 hospitais com melhor desempenho, somente 7 são comuns no estadio 1, 6 no estadio 2 e 5 no estadio 3. Neste grupo o hospital com melhor desempenho é o número 41, com excepção nos doentes mais graves, sendo nesta situação o hospital 13 o que apresenta melhor desempenho.

Para os 10 hospitais com pior desempenho, 7 hospitais são comuns no estadio 1, 8 no estadio 2 e 6 no estadio 3. O hospital 9 é o menos eficiente, com excepção nos doentes mais graves, na qual o hospital 61 é o menos eficiente.

Ainda em relação à associação entre o desempenho global dos hospitais e a gravidade dos doentes, observa-se que o coeficiente de correlação de Pearson é de 0.964, entre todos os doentes e os menos graves, de 0,919 entre todos os doentes e os doentes com severidade intermédia e de 0.686 em relação aos doentes mais graves.

A manutenção desta associação e com maior intensidade nos doentes menos graves, reforça os comentários feitos anteriormente, ou seja, que a maior eficiência dos hospitais parece estar associada com uma predominância de doentes menos graves.

Para o total da produção dos hospitais observa-se que nos 10 mais eficientes estão 7 hospitais do Tipo III e 3 do Tipo I. Na distribuição Regional, 6 encontram-se na Região IV e 2 nas regiões III e V.

Para os 10 hospitais menos eficientes, 7 são do Tipo III e 3 do Tipo I, enquanto que a distribuição regional é mais dispersa, com 4 hospitais a pertencerem à Região III, 3 à IV, 2 à V e 1 à Região I.

Estes resultados parecem mais uma vez evidenciar que a pesquisa e identificação de critérios para a avaliação e comparação do desempenho dos hospitais predominantemente baseados em critérios de oferta não são consistentes, nem aconselháveis.

Por outro lado, o panorama a nível regional parece evidenciar que, tanto a administração a nível central e regional, como os mecanismos de regulação não têm sido efectivos, nem gerado incentivos consistentes, visto que a eficiência parece estar muito mais associada ao desempenho próprio de cada uma das organizações de saúde.

Para os doentes menos graves, nos 10 hospitais mais eficientes, 6 são do Tipo III, 3 do Tipo I e 1 do Tipo II, estando 5 na Região IV, 4 na V e 1 na III. Para os 10 hospitais com pior desempenho, 6 são do Tipo III e 4 do I, enquanto que 4 estão localizados na Região III, 2 nas III e 1 nas I e II.

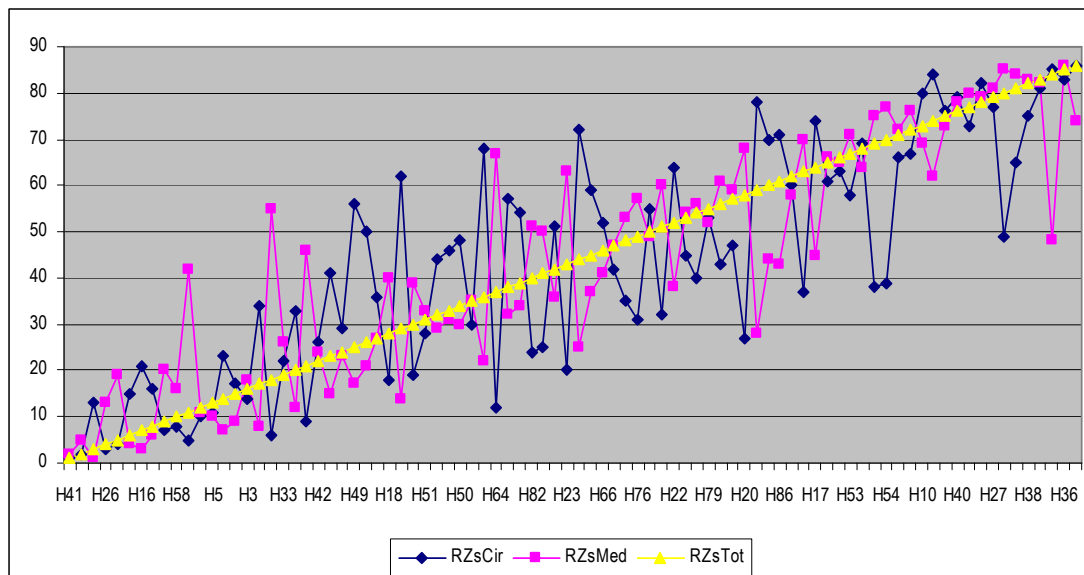
Para o escalão de gravidade intermédia, nos 10 hospitais com melhor eficiência 5 são do Tipo I e 5 do Tipo III, com 4 a pertencerem à Região IV e V e 2 à Região III. Com pior eficiência, encontram-se 6 hospitais do Tipo III e 4 do Tipo I, com 3 hospitais a serem encontrados nas Regiões III, IV, e V e o restante na I.

Finalmente para os doentes mais graves, nos hospitais mais eficientes encontram-se igualmente 5 hospitais do Tipo I e do Tipo III, estando 4 na Região IV e 3 nas Regiões III e V. Para os hospitais menos eficientes 8 são do Tipo III e 2 do Tipo I, estando 3 localizados nas Regiões III e IV, 3 na II e 2 nas Regiões I e V.

Como comentário final deve referir-se que embora para todos os estadios de gravidade exista um grande predomínio dos hospitais do Tipo I e do Tipo III, tanto para as organizações mais eficientes, como para as menos eficientes, observa-se ainda que, muitas das vezes não são os mesmos hospitais a apresentarem melhores ou piores níveis de eficiência. Assim, esta análise mais desagregada vem igualmente confirmar o referido, essencialmente no que se refere à insuficiência e limitações dos mecanismos da oferta para se compararem hospitais e à necessidade de um reforço da gestão e da regulação das organizações de saúde.

De seguida passa-se a analisar a eficiência dos hospitais para os casos médicos e cirúrgicos (ver Figura 47). No Anexo XI podem ser consultados os valores.

Figura 47
Ordenação dos Hospitais, Eficiência
Casos Médicos e Cirúrgicos e Total dos Doentes



O desempenho dos hospitais entre casos cirúrgicos e médicos é diferente, sendo o *K de Cohen* de 0.212. A relação entre casos médicos e total de doentes também apresenta uma fraca concordância no desempenho dos hospitais com o *K* a ser de 0.024. No que se refere aos casos cirúrgicos o panorama mantém-se, apresentando esta estatística um valor de 0.024.

É ainda de referir que a demora média global está mais associada com a dos casos médicos, do que com a dos casos cirúrgicos, sendo o coeficiente de correlação de Pearson de respectivamente 0.897 e de 0.840.

Para os casos cirúrgicos, entre os 10 hospitais com melhor desempenho 7 são do Tipo III e 1 dos Tipos I, II e IV, encontrando-se 6 na Região V, e na IV e 1 na Região III. Destes 10 hospitais somente 2 apresentam igualmente o melhor desempenho para os casos médicos, subindo este valor para 5 hospitais quando se considera a totalidade dos doentes tratados.

Nos 10 hospitais com pior desempenho nos casos cirúrgicos 6 são do Tipo III e 4 do Tipo I, estando 4 localizados na Região III, 3 na V, 2 na IV e 1 na Região II. Destes 10 hospitais 5 estão igualmente entre os que apresentam pior desempenho nos casos médicos e 6 quando se considera o total da produção dos hospitais.

Nos casos médicos, para os hospitais com melhor desempenho 5 são do Tipo I e 5 do Tipo III, estando 7 localizados na Região IV, 2 na III e 1 na Região V. Destes 10 hospitais somente 2 ocupam a mesma posição nos casos cirúrgicos, aumentando este número para 6 quando se considera toda a produção.

Ainda para os casos médicos, nos hospitais com pior desempenho 9 são do Tipo III e 1 do Tipo I, encontrando-se 3 nas Regiões III e V, 2 na Região IV e 1 nas Regiões I e II. Destes hospitais 5 são comuns quando se considera o desempenho nos casos cirúrgicos, passando este valor para 8 hospitais quando se consideram todos os doentes tratados.

Para se analisarem as doenças serão utilizados, mais uma vez, os Grandes Agrupamentos de Doenças (GADs), sendo identificados os cinco hospitais mais e menos eficientes em função do método directo (ver Quadro XLV).

Quadro XLV
Hospitais Mais e Menos Eficientes por Grandes Agrupamentos de Doenças

Grandes Agrupamentos de Doenças	Hospitais Mais Eficientes	Hospitais Menos Eficientes
Doenças da mama	16, 25, 33, 41 e 20	86, 87, 10, 88 e 1
Doenças do sistema nervoso central	13, 10, 58, 16 e 25	4, 44, 29, 9 e 43
Doenças cardiovasculares e do coração	8, 5, 9, 55 e 3	14, 1, 27, 29 e 36
Doenças dermatológicas	26, 9, 1, 41 e 2	10, 60, 12, 86 e 19
Doenças do ouvido, nariz e garganta	21, 41, 48, 35 e 58	8, 12, 19, 44 e 10
Doenças gastrointestinais	26, 3, 41, 25 e 1	10, 4, 9, 60 e 29
Doenças ginecológicas	2, 18, 48, 19 e 20	43, 61, 46, 9 e 88
Doenças hepatobiliares	6, 1, 26, 41 e 3	19, 10, 12, 2 e 36
Doenças hematológicas	4, 2, 9, 6 e 35	44, 88, 29, 38 e 7
Doenças Imunológicas	16, 7, 8, 10 e 1	15, 25, 25, 4 e 40
Doenças linfáticas	2, 87, 1, 35 e 33	9, 13, 44, 55 e 29
Doenças dos órgãos genitais masculinos	4, 21, 26, 41 e 8	27, 9, 36, 10 e 1
Doenças musculoesqueléticas	25, 41, 9, 1 e 13	9, 12, 40, 17 e 38
Doenças endócrinas e metabólicas	23, 13, 26, 3 e 32	1, 7, 36, 12 e 27
Doenças obstétricas	26, 39, 85, 41 e 56	9, 20, 61, 46 e 27
Doenças oftalmológicas	1, 34, 9, 41 e 48	4, 38, 12, 10 e 61
Doenças do sistema nervoso periférico	26, 6, 22, 71 e 55	9, 43, 4, 1 e 31
Doenças psiquiátricas	21, 26, 31, 55 e 1	3, 9, 10, 60, e 64
Doenças renais	5, 7, 4, 88 e 15	9, 19, 29, 61 e 27
Doenças respiratórias	6, 2, 25, 32 e 41	9, 44, 19, 43 e 15
Doenças multisistémicas	87, 13, 9, 32 e 3	88, 86, 12, 29 e 44
Doenças vasculares (excepto coração)	1, 5, 9, 13 e 25	14, 2, 8, 29 e 36
Doenças não especificadas	9, 87, 84, 58 e 16	78, 27, 45, 68 e 60
Outras doenças	15, 87, 3, 18 e 86	88, 60, 44, 10 e 27

Para os hospitais mais eficientes observa-se que, para os 24 Grandes Agrupamentos de Doenças, existem 14 hospitais diferentes. Destes o hospital 26 é o mais eficiente em 4 doenças, enquanto que os hospitais 1, 2, 4, 6, 16, 21 e 25 são mais eficientes em duas doenças.

Quando se consideram os cinco hospitais mais eficientes para cada uma das doenças, verifica-se que para as 120 permutações possíveis, existem 38 hospitais que aparecem citados pelo menos uma vez. Os hospitais que apresentam maior frequência são o 41 (10 vezes), o 1 (9 vezes), o 26 (8 vezes) e os 3, 9 e 25 (7 vezes cada um).

Na situação contrária, para os hospitais menos eficientes, existem 13 organizações distintas citadas, das quais somente 5 mais que uma vez. São os hospitais 9 (7 vezes) e 4, 10, 14 e 88 (2 vezes).

Considerando agora os 5 hospitais menos eficientes, existem 37 organizações com citações. Os hospitais que aparecem em maior número de doenças com pior desempenho são os hospitais 9 (12 vezes), 10 e 29 (8 vezes), 27 (7 vezes) e 12 e 44 (6 vezes).

À semelhança do referido para a efectividade observa-se que existem alguns hospitais que surgem referenciados com melhor e pior eficiência em função da respectiva doença tratada. Os melhores exemplos para esta situação são encontrados nos hospitais 1 (9 vezes entre os 5 hospitais mais eficientes e 5 vezes entre os e menos eficientes), 4 (3 vezes entre os 5 hospitais mais eficientes e 5 vezes entre os 5 menos eficientes) e 9 (7 vezes entre os 5 hospitais mais eficientes e 12 vezes entre os 5 hospitais menos eficientes).

Estes resultados permitem afirmar que o desempenho dos hospitais é variável em função da eficiência, dependendo esta, por sua vez, do tipo de doença que é tratado, o que mais uma vez vem confirmar a necessidade de se reforçarem os mecanismos de regulação e de administração nos hospitais.

Outra questão interessante a discutir será a necessidade ou não de se utilizarem estes valores desagregados por doença, para futuras actividades de planeamento e de afectação de recursos, visto que os resultados serão forçosamente diferentes daqueles que utilizam valores mais agregados.

5.2.3. Desempenho Global

Atento o estado da arte da avaliação do desempenho dos hospitais é reconhecido que três componentes assumem particular relevo – a estrutura, o processo e os resultados.

Contudo, este estudo visa essencialmente a avaliação dos resultados em saúde, embora também se pudessem incorporar outros aspectos, essencialmente os relacionados com a produção, nomeadamente a medição da produção (Índice de Casemix) e o perfil das admissões (concentração da produção, complexidade e severidade).

A principal razão justificativa da escolha exclusiva de variáveis explicativas dos resultados deriva do facto de os indicadores de produção e do perfil das admissões reflectirem algumas questões ligadas com a estrutura (limitações na oferta de recursos e de serviços dos hospitais), ou de processo (o actual estado da arte num determinado hospital não permite o tratamento de determinadas situações de saúde), ou ainda de práticas explícitas e criteriosas de hospitais, de forma a optimizarem os seus resultados.

Nesta conformidade, estes dois aspectos – medição da produção e perfil das admissões – são considerados como “inputs” ou ainda restrições dos hospitais, as quais por sua vez podem condicionar os resultados de saúde.

Assim, escolhem-se somente os indicadores de efectividade e de eficiência, restando assim discutir qual o peso específico de cada um no modelo global de avaliação do desempenho.

Esta questão não é puramente académica, visto que a discussão central é a seguinte:

Qual o objectivo a optimizar?

A efectividade?

A efectividade, com restrições na eficiência?

A eficiência?

A eficiência, com restrições na efectividade?

Sem se pretender dar respostas definitivas, pode remeter-se a resposta a estas questões com uma nova interrogação?

Quais as preferências que contam?

No actual momento do sector da saúde em Portugal é natural esperar que, se as preferências a privilegiar forem dos proprietários ou dos gestores que o principal objectivo seja optimizar a eficiência, quanto muito limitada a um nível razoável/aceitável de qualidade.

Se, por outro lado, as preferências a considerar forem dos consumidores, ou mesmo dos prestadores, pode afirmar-se que, embora por razões distintas, o principal objectivo seja otimizar a efectividade, igualmente quanto muito limitada por um nível razoável/aceitável de eficiência.

Em termos mais operacionais, interessa igualmente estudar se a avaliação do desempenho dos hospitais resulta completamente diferente quando se utiliza cada uma das perspectivas – efectividade e eficiência – quando comparadas com um indicador conjunto.

Para a criação deste indicador conjunto foram estabelecidas três alternativas:

- Valorização idêntica da efectividade e da eficiência;
- Maior valorização da efectividade;
- Maior valorização da eficiência.

Por outro lado, deve ainda ter-se em consideração a metodologia para se valorizarem os indicadores de efectividade e de eficiência. A este propósito duas perspectivas se colocam:

- Utilização das ordenações dos hospitais para a efectividade e para a eficiência;
- Utilização directa dos valores da efectividade e da eficiência de cada hospital.

Para a primeira perspectiva, as principais razões justificativas para a sua utilização são encontradas na facilidade de cálculo e na percepção do indicador.

Para a segunda perspectiva, para além das razões referidas anteriormente, deve ser acrescido o facto de incluir a dimensão do respectivo desempenho, o qual aparece mais esbatido quando se utiliza a ordenação dos hospitais.

Neste estudo, para se caracterizar as realidades decorrentes da utilização de diferentes metodologias para se avaliar o desempenho global dos hospitais, irão ser construídos alguns cenários.

Quando se utilizam os valores da ordenação dos hospitais decorrentes da utilização do método indirecto, as alternativas são as seguintes:

- Cenário A - valorização idêntica da efectividade e da eficiência;
 $\text{Indicador conjunto} = (\text{Ordenação da efectividade} + \text{Ordenação da eficiência}) / 2 \text{ (RGA)};$
- Cenário B – maior valorização da efectividade
 $\text{Indicador conjunto} = (2 * \text{Ordenação da efectividade} + \text{Ordenação da eficiência}) / 3 \text{ (RGB)};$

- Cenário C – maior valorização da eficiência
Indicador conjunto = (Ordenação da efectividade + 2 * Ordenação da eficiência) / 3 (RGC).

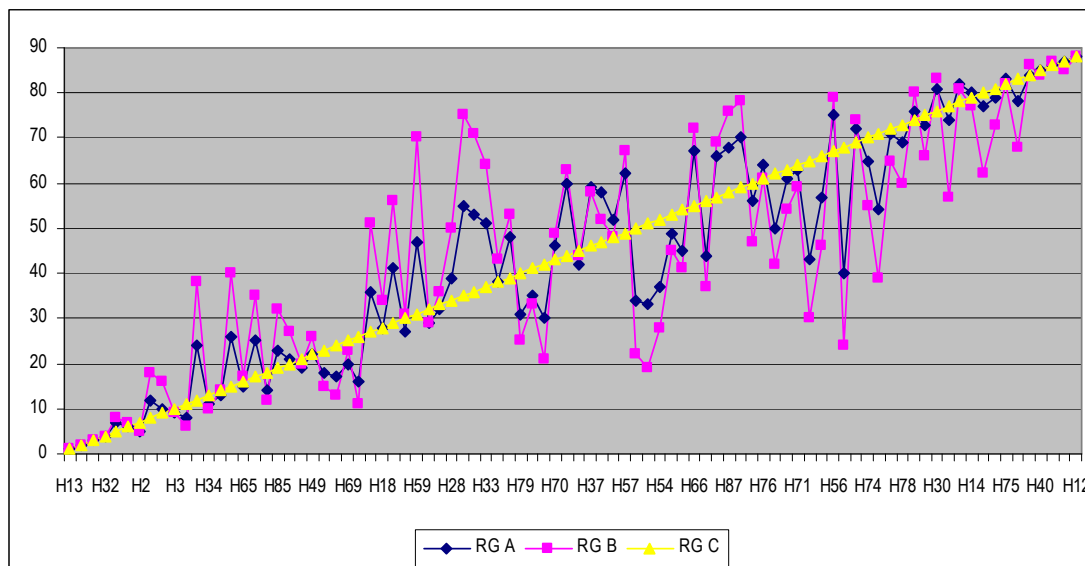
Quando se utilizam os indicadores de efectividade e da eficiência decorrentes do método directo, as alternativas são as seguintes:

- Cenário I – valorização idêntica da efectividade e da eficiência
Indicador conjunto = (“z score” da efectividade + “z score” da eficiência) / 2 (RGI);
- Cenário II – maior valorização da efectividade
Indicador conjunto = (2 * “z score” da efectividade + “z score” da eficiência) / 3 (RGII);
- Cenário III – maior valorização da eficiência
Indicador conjunto = (“z score” da efectividade + 2 * “z score” da eficiência) / 3 (RGIII).

Num primeiro momento serão feitas comparações entre as diferentes formas de valorizar o indicador conjunto de desempenho hospitalar, tanto quando se utilizam os valores das ordenações para a efectividade e para a eficiência, como quando se utilizam directamente os indicadores de efectividade e de eficiência.

Na Figura 48 são apresentados os indicadores conjuntos do desempenho dos hospitais para os Cenários A, B e C. No Anexo XII podem ser consultados os valores.

Figura 48
Desempenho Global dos Hospitais
Ordenações de Efectividade e de Eficiência

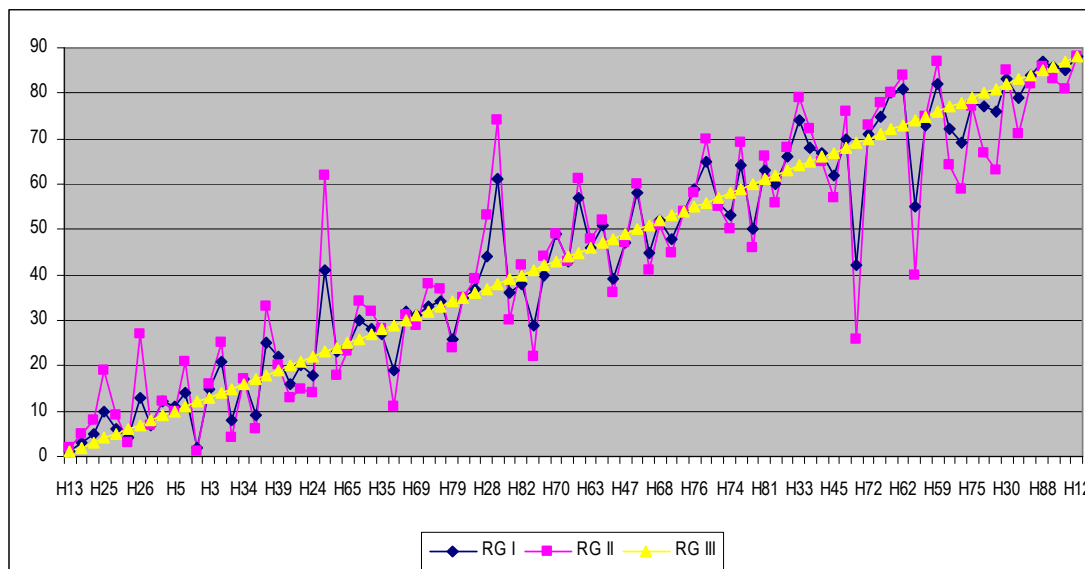


Conforme se verifica os três indicadores estão a dar informações diferentes:

- O *K de Cohen* é de 0.080 entre o indicador que dá valorização idêntica à efectividade e à eficiência (Cenário A) e aquele que decorre de uma maior valorização da efectividade (Cenário B). Esta estatística é de 0.138, quando se comparam os indicadores dos Cenários A e C (maior valorização da eficiência). Por sua vez, o *K de Cohen* é de 0.092 entre os indicadores dos Cenários B e C;
- Os hospitais 13, 41, 48, 32 e 2 são os melhores no Cenário A, enquanto que os hospitais 12, 29, 43, 40 e 88 são os piores igualmente nesta perspectiva de avaliação;
- Para o Cenário B os hospitais 13, 41, 48, 32 e 2 são os melhores. Por sua vez os hospitais 12, 43, 88, 29 e 40 são os que apresentam pior desempenho;
- No Cenário C os hospitais 13, 41, 48, 32 e 58 são os melhores, enquanto que os hospitais 12, 29, 43, 40 e 88 são os piores.

Na Figura 49 são apresentados os indicadores conjuntos do desempenho dos hospitais para os Cenários I, II e III. No Anexo XII podem ser consultados os valores.

Figura 49
Desempenho Global dos Hospitais
“Z score” de Efectividade e de Eficiência



Tendo em atenção a estatística *K de Cohen* os três indicadores estão a dar informações distintas, sendo os seus valores de 0.103, de 0.092 e de 0.057, respectivamente entre os Cenários I e II, I e III e II e III.

Para uma valorização idêntica da efectividade e da eficiência os melhores hospitais são os números 13, 9, 41, 2 e 48 e os hospitais 12, 88, 43, 29 e 40 são os piores.

Dando maior peso à efectividade são os hospitais 9, 13, 2, 61 e 41 os que apresentam melhor desempenho, enquanto que os hospitais 12, 59, 88, 30 e 62 são os piores.

Quando se valoriza mais a eficiência surgem os hospitais 13, 41, 48, 25 e 32 com melhor desempenho. Por sua vez os hospitais 12, 29, 43, 88 e 40 são os que apresentam pior desempenho.

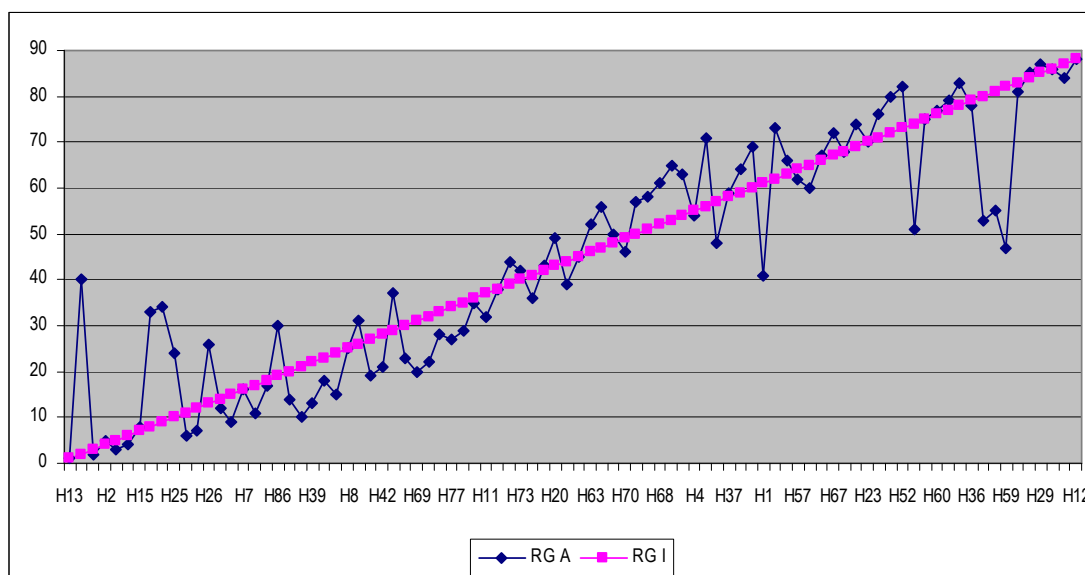
Estes resultados provam que o desempenho dos hospitais depende fortemente da perspectiva que se pretende valorizar, visto que a sua ordenação é diferente quando se ponderam de formas distintas a efectividade e a eficiência.

Assim, na ausência de outro tipo de estudos, essencialmente aqueles que visam investigar qual a dimensão que deve ser privilegiada para se avaliar o desempenho dos hospitais, sugere-se que o indicador conjunto a utilizar seja aquele que dá a mesma importância à efectividade e à eficiência.

Neste sentido, resta agora discutir qual o método que deve servir de base para esta mesma avaliação: o baseado nas ordenações (RGA)? Ou o baseado nos “z scores” (RGI)?

Na Figura 50 apresentam-se as ordenações dos hospitais baseados nestes dois critérios. No Anexo XII podem ser consultados os respectivos valores.

Figura 50
Desempenho Global dos Hospitais
“Z score” de Efectividade e de Eficiência (RGI) e
Ordenações de Efectividade e de Eficiência (RGA)



Mais uma vez se observa que estes dois sistemas de classificação estão a produzir resultados completamente diferentes, sendo o *K de Cohen* de 0.103.

Nos 10 hospitais com melhor desempenho quando se utiliza o método directo, somente 6 ocupam iguais posições quando se utiliza as respectivas ordenações. Mais relevante é a situação encontrada no hospital 9 o qual é o segundo melhor em função do “z score” e apenas o 40º na classificação baseada nas ordenações. Comportamentos similares são encontrados nos hospitais 10, 25 e 61 (consultar Anexo XII).

Para os 10 hospitais com pior desempenho segundo o método directo encontram-se igualmente 6 quando se utilizam as respectivas ordenações. Os hospitais 55, 59 e 62 são os que apresentam comportamentos mais heterogéneos entre os dois sistemas de classificação (consultar Anexo XII).

Para os 88 hospitais verifica-se que as maiores diferenças entre os dois sistemas de classificação são encontradas nos hospitais 9, 61, 10, 31 e 25 quando a ordenação proporcionada pelo “z score” é inferior à baseada nas ordenações e nos hospitais 59, 55, 62, 33 e 1 quando a situação é exactamente a contrária, ou seja ordenação dos “z score” superior à observada nas ordenações.

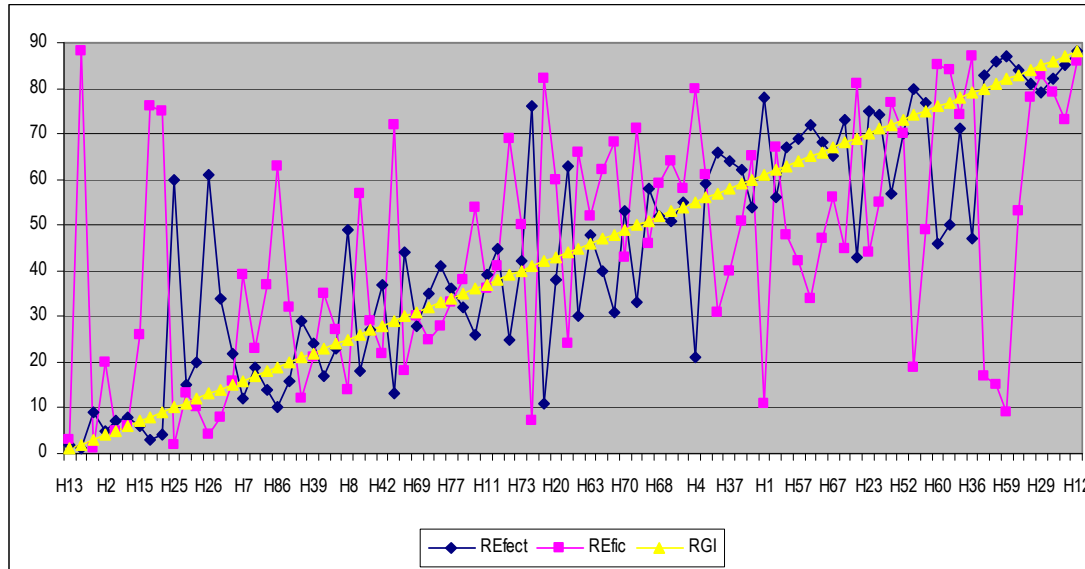
Tendo em atenção estes resultados sugere-se que seja utilizada a metodologia que se baseia nos valores dos indicadores de efectividade e de eficiência (“z score”), essencialmente porque desta forma se discrimina melhor os hospitais que apresentam melhores resultados somente numa das perspectivas em análise.

Assim, de acordo com as metodologias sugeridas, na Figura 51 é apresentado o desempenho global dos hospitais (RGI), bem como as respectivas posições quando se considera somente a efectividade (REfect) e a eficiência (REfic). No Anexo XII podem ser consultados os respectivos valores.

Mais uma vez se observa que o desempenho global dos hospitais é completamente diferente do observado, tanto para a efectividade, como para a eficiência, sendo o *K de Cohen* de respectivamente 0.023 e de 0.011. Este resultado é função directa das diferenças encontradas entre a efectividade e a eficiência dos hospitais, dado que esta mesma estatística é de 0.000.

A análise de correlação através do coeficiente de Pearson, sendo esta estatística de 0.856 e de 0.498, respectivamente entre o indicador global e o da efectividade e entre o indicador global e o de eficiência, parece evidenciar que o desempenho global dos hospitais está mais associado com a efectividade dos cuidados prestados, do que com a eficiência.

Figura 51
Ordenação dos Hospitais
Desempenho Global, Efectividade e Eficiência



Tendo em atenção estes elementos pode afirmar-se que o hospital com melhor desempenho em Portugal é o número 13, o qual ocupa a 2ª posição na efectividade e a 3ª na eficiência. O hospital com pior desempenho, o número 12, ocupa igualmente a última posição na efectividade e a antepenúltima na eficiência.

Para os 10 hospitais com melhor desempenho global devem fazer-se as seguintes observações:

- 9 hospitais são comuns à classificação da efectividade. O hospital 25 é o que entra nesta ordenação global, ocupando a 60ª posição em termos de efectividade. Por sua vez, o hospital 86 ocupava a 10ª posição na efectividade, apenas está classificado na 19ª posição em termos globais;
- No que se refere à eficiência existem 5 hospitais comuns. Os hospitais 9, 61, e 10 são os que apresentam comportamentos mais distintos, visto que ocupam respectivamente as 88ª (última), 76ª e 75ª posições na eficiência. No que se refere aos hospitais que se encontram bem posicionados na eficiência e que ocupam posições menos relevantes no desempenho global, podem referir-se o número 16 (7º na eficiência e 41º no desempenho global) e o número 59 (9º na eficiência e 82º no desempenho global);
- Existem 5 hospitais dos Tipos III e I, estando 5 localizados na Região de Saúde III, 4 na IV e 1 na Região V.

Para os 10 piores hospitais, 9 são comuns à ordenação da efectividade e 6 à da eficiência.

Com maiores discrepâncias no que se refere à efectividade encontra-se o hospital 36, visto que ocupa a 79^a posição no desempenho global e a 47^a nesta dimensão.

No que se refere à eficiência, para além do referido para o hospital 59, deve ainda evidenciar-se a situação dos hospitais 55 e 62, visto que ocupam a 17^a e a 15^a posições entre os mais eficientes e a 80^a e 81^a posições no desempenho global.

Destes 10 hospitais 8 são do Tipo III e 1 dos Tipos I e V, estando 5 localizados na Região de Saúde IV, 2 na III e 1 nas Regiões I, II e IV.

Estes resultados vêm mais uma vez confirmar que os critérios exclusivamente baseados na oferta, só por si, não justificam o desempenho dos hospitais, nem constituem uma base sólida para se construírem bases de comparação.

Bons exemplos ilustrativos são os factos de o melhor e o pior hospital em Portugal pertencerem ao mesmo tipo e estarem localizados na mesma Região de Saúde e ainda de os hospitais do Tipo III serem predominantes, tanto entre aqueles que apresentam o melhor, como o pior desempenho.

Em seguida, à semelhança do efectuado para a efectividade e para a eficiência serão avaliados os hospitais por tipo de tratamento (cirúrgico e médico) e por doença (Grandes Agrupamentos de Doenças).

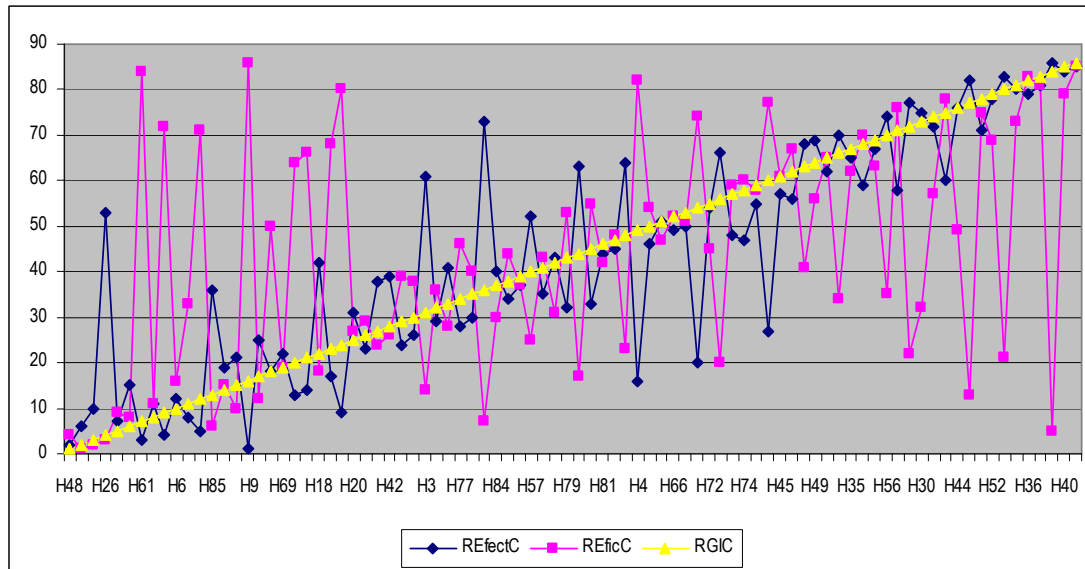
Para qualquer destas perspectivas de análise serão somente utilizados os indicadores de efectividade e da eficiência resultantes do método directo ("z score").

Na Figura 52 é apresentado o desempenho global dos hospitais (RGIC) para os casos cirúrgicos, bem como as respectivas posições quando se considera somente a efectividade (REfectC) e a eficiência (REficC). No Anexo XII podem ser consultados os respectivos valores.

Não existe concordância entre a ordenação dos hospitais para o desempenho global e com a efectividade e eficiência, sendo o *K de Cohen* respectivamente de 0.024 e 0.035.

Contudo, o desempenho global parece mais uma vez estar mais associado com a efectividade do que com a eficiência, visto que o coeficiente de correlação de Pearson é de respectivamente de 0.881 e de 0.453.

Figura 52
Ordenação dos Hospitais
Desempenho Global, Efectividade e Eficiência
Casos Cirúrgicos



Para os casos cirúrgicos o melhor hospital é o número 48, o qual ocupa a 2ª posição na efectividade e a 4ª na eficiência. O hospital com pior desempenho é o número 12, o qual ocupa as penúltimas posições, respectivamente para a efectividade e para a eficiência.

Para os 10 hospitais com melhor desempenho global, embora sendo 6 comuns tanto com a ordenação da efectividade como com a da eficiência, deve referir-se que não são as mesmas organizações de saúde.

Destes 10 hospitais, 7 são do Tipo III, 2 do Tipo I e 1 do Tipo V, estando 5 localizados na Região de Saúde IV, 3 na Região V e 2 na Região III.

Para os 10 hospitais com pior desempenho global, 9 são comuns com a ordenação da efectividade e 4 com o da eficiência.

Destes 10 hospitais, 6 são do Tipo III e 4 do Tipo I, estando 3 localizados na Região IV, 2 nas Regiões I, III e V e 1 na Região II.

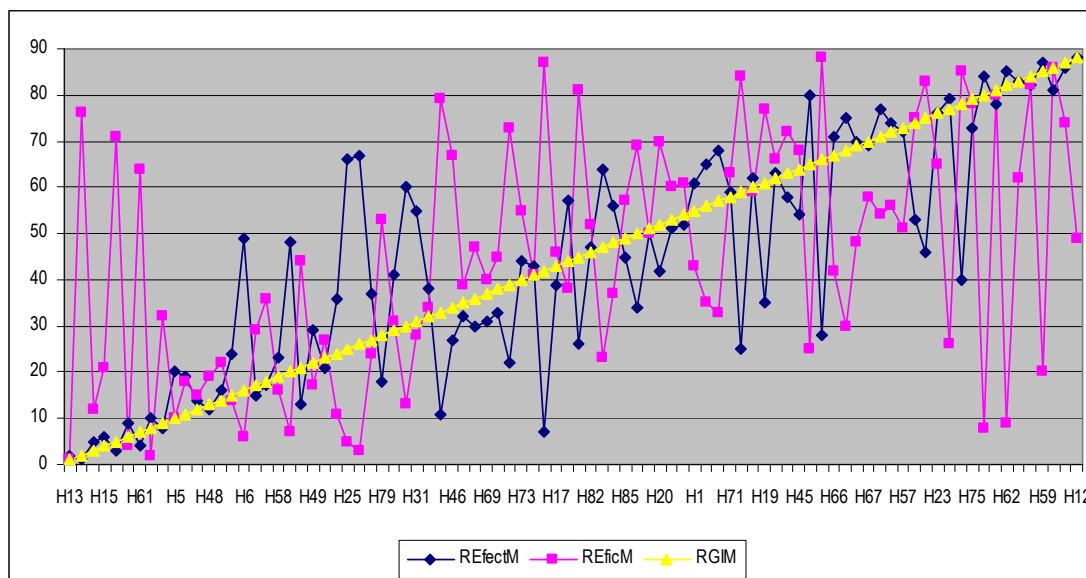
Comparando com as ordenações da efectividade, os hospitais que apresentaram melhorias mais substanciais no desempenho global são os números 26, 59, 3, 85 e 8 enquanto que aqueles em que se observou uma maior descida na ordenação global são os hospitais 17, 27, 4, 31 e 10.

A comparação com a eficiência permite afirmar que os hospitais com melhoria na ordenação global são os números 61, 9, 87, 86 e 10. Deve ainda referir-se que para estes hospitais existem comportamentos bastante diferentes entre efectividade e eficiência, ocupando quase sempre as primeiras posições para a primeira dimensão e as últimas para a eficiência.

Na situação contrária encontram-se os hospitais 1, 13, 16, 33 e 30. Destes hospitais deve ainda evidenciar-se as situações do número 1, o qual é o 5º mais eficiente e o menos efectivo e ainda a do hospital 13, visto que é o 13º mais eficiente e o 5º menos efectivo.

Na Figura 53 é apresentado o desempenho global dos hospitais (RGIM) para os casos médicos, bem como as respectivas posições quando se considera somente a efectividade (REfectM) e a eficiência (REficM). No Anexo XII podem ser consultados os respectivos valores.

Figura 53
Ordenação dos Hospitais
Desempenho Global, Efectividade e Eficiência
Casos Médicos



Para os casos médicos não existe igualmente concordância na ordenação dos hospitais. O *K de Cohen* é de 0.023 entre o desempenho global e a efectividade e de 0.034 entre o desempenho global e a eficiência. Pode ainda referir-se que entre a efectividade e a eficiência não existe igualmente concordância na ordenação dos hospitais, sendo o *K de Cohen* de 0.011.

Por outro lado, os resultados globais estão mais associados com a efectividade dos cuidados prestados do que com a eficiência, sendo o coeficiente de correlação de Pearson respectivamente de 0.842 e de 0.516.

O hospital 13 aparece com o melhor desempenho global, ocupando a 2ª posição na efectividade e a 1ª na eficiência, enquanto que o hospital 12 apresenta o pior desempenho global, situação que mantém na efectividade, embora ocupe uma posição intermédia na eficiência (49ª).

Para os 10 hospitais com melhor desempenho, 8 são comuns na ordenação com a efectividade e somente 4 com o desempenho em termos de eficiência.

Destes 10 hospitais, 7 são do Tipo I e 3 do Tipo III, estando 5 localizados na Região III, 4 na IV e 1 na Região V.

Para os 10 hospitais com pior desempenho, 8 são comuns com a ordenação da efectividade e 3 com o da eficiência.

Destes 10 hospitais, 7 são do Tipo III e 1 dos Tipos I, IV e V, estando 6 localizados na Região IV e 1 em cada uma das restantes Regiões de Saúde.

A comparação com a efectividade permite evidenciar a situação dos hospitais 16, 25, 6, 26 e 8, na qual o desempenho global é bastante superior ao da efectividade e a dos hospitais 38, 36, 44, 60 e 27 na qual a situação é exactamente a contrária. Os hospitais 12, 23 e 30 apresentam um desempenho exactamente igual para estes dois atributos.

Tendo como referência a eficiência deve referir-se que os hospitais 9, 10, 61, 54 e 44 apresentam um desempenho global superior ao obtido na eficiência, enquanto que nos hospitais 62, 55, 59, 33 e 87 a situação é exactamente a contrária. Por sua vez, os hospitais 5, 13, 18 e 29 apresentam desempenhos iguais para estes dois atributos.

Em seguida serão apresentados os hospitais com melhor e pior desempenho global pelo Grande Agrupamento de Doenças do Disease Staging. À semelhança do realizado para a efectividade e para a eficiência, somente serão identificados os 5 primeiros e últimos por doença (ver Quadro XLVI).

Existem 39 hospitais que não aparecem nenhuma vez referenciados entre aqueles que apresentam melhor ou pior desempenho global.

Para os hospitais com melhor desempenho existem 34 organizações de saúde que aparecem pelo menos uma vez referenciadas entre as 5 melhores. Por outro lado, observa-se a existência de 13 hospitais diferentes para o melhor desempenho global.

Quadro XLVI
Hospitais com Melhor e Pior Desempenho Global
por Grandes Agrupamentos de Doenças

Grandes Agrupamentos de Doenças	Hospitais com Melhor Desempenho	Hospitais com Pior Desempenho
Doenças da mama	13, 2, 9, 3 e 55	88, 87, 86, 29 e 62
Doenças do sistema nervoso central	13, 10, 61, 32 e 58	12, 1, 43, 29 e 30
Doenças cardiovasculares e do coração	9, 10, 31, 42 e 5	12, 59, 30, 33 e 29
Doenças dermatológicas	26, 4, 61, 65 e 9	12, 36, 13, 83 e 40
Doenças do ouvido, nariz e garganta	9, 86, 13, 2 e 48	88, 87, 8, 12 e 60
Doenças gastrointestinais	9, 13, 26, 15 e 41	12, 44, 55, 88 e 59
Doenças ginecológicas	2, 48, 18, 20 e 19	88, 87, 46, 36 e 1
Doenças hepatobiliares	9, 13, 1, 6 e 26	40, 55, 12, 59 e 60
Doenças hematológicas	4, 9, 26, 15 e 16	88, 87, 23, 59 e 38
Doenças imunológicas	9, 7, 2, 1, 11	26, 25, 10, 12 e 48
Doenças linfáticas	35, 10, 1, 3 e 6	88, 55, 87, 29 e 86
Doenças dos órgãos genitais masculinos	3, 34, 13, 48 e 6	88, 87, 16, 56 e 57
Doenças musculoesqueléticas	25, 41, 59, 48 e 8	12, 1, 9, 13 e 40
Doenças endócrinas e metabólicas	9, 13, 15, 26 e 10	62, 59, 30, 12 e 33
Doenças obstétricas	20, 41, 8, 54 e 55	10, 61, 1, 29 e 39
Doenças oftalmológicas	10, 13, 1, 34, 2	7, 40, 3, 30 e 26
Doenças do sistema nervoso periférico	9, 2, 3, 7 e 22	43, 4, 54, 40 e 29
Doenças psiquiátricas	13, 15, 1, 26 e 7	62, 3, 59, 10 e 28
Doenças renais	5, 34, 25, 2 e 15	33, 59, 27, 55 e 88
Doenças respiratórias	2, 6, 61, 25 e 41	9, 44, 43, 62, 15
Doenças multisistêmicas	13, 87, 9, 2 e 25	88, 12, 40, 29 e 33
Doenças vasculares (excepto coração)	5, 9, 10, 13 e 26	12, 4, 14, 2 e 60
Doenças não especificadas	9, 87, 58, 84 e 35	26, 78, 79, 45 e 66
Outras doenças	87, 15, 86, 13 e 41	88, 38, 1, 40 e 10

Quando se analisam os hospitais com pior desempenho global, existem 45 organizações de saúde diferentes com pelo menos uma referência entre os 5 piores, enquanto que para o pior desempenho este número desce para 11.

Deve ainda referir-se que para os 59 hospitais identificados entre aqueles que apresentam melhor ou pior desempenho, 19 organizações de saúde recebem somente uma citação. Entre estes últimos hospitais pode evidenciar-se que 13 correspondem a situações identificadas para o pior nível de desempenho global.

Ainda para os 59 hospitais identificados entre aqueles que apresentam melhor ou pior desempenho, deve referir-se que 14 correspondem a situações identificadas com o melhor desempenho e 25 somente com pior desempenho.

Os restantes 20 hospitais são classificados entre aqueles que simultaneamente apresentam bom e mau desempenho.

Analisando com maior detalhe os elementos do Quadro XLVI podem fazer-se os seguintes comentários:

- Para o melhor desempenho, dos 13 hospitais identificados, 9 aparecem citados somente uma vez. Para os restantes 4 hospitais, o número 9 recebe 8 citações, o número 13 surge 4 vezes e os números 2 e 5 aparecem 2 vezes;
- Para os 34 hospitais identificados entre aqueles que surgem entre os cinco melhores no desempenho global, o número 9 recebe 13 citações, enquanto que os números 13, 2 e 26 aparecem referenciados respectivamente, 12, 9 e 7 vezes;
- Para os 11 hospitais com pior desempenho, 7 são identificados somente uma vez, enquanto que os números 88 e 12 recebem 7 e 6 citações, respectivamente;
- Quando se consideram os cinco hospitais com pior desempenho, os números 12 e 88 recebem 11 e 10 citações, enquanto que os números 29, 40 e 59 são identificados 7 vezes;
- Deve ainda referir-se que o hospital 1 é aquele que apresenta um comportamento mais irregular, visto que aparece 5 vezes citado, tanto para o melhor como para o pior desempenho global (o desempenho global é avaliado para as cinco primeiras e para as cinco últimas posições). A este respeito merece ainda destaque a situação dos hospitais 10 (6 citações para o melhor e 4 para o pior desempenho), 26 (7 e 3 citações para a mesma ordem) e 87 (3 e 6 citações para o melhor e pior desempenho).

Estes aspectos confirmam o que foi referido para a avaliação da efectividade e da eficiência, em que por vezes as diferenças intra-hospitalares são quase tão relevantes como as existentes entre aquelas organizações de saúde, pelo que se reitera a necessidade de se reforçarem os mecanismos de regulação e de administração do sector hospitalar português.

5.3. Síntese dos Resultados

Em seguida serão apresentados alguns Quadros resumo das dimensões da produção e do desempenho dos hospitais e dos respectivos indicadores.

As dimensões em estudo são:

- Produção;
- Perfil das Admissões;
- Efectividade;
- Eficiência;
- Desempenho Global.

Para estas dimensões serão somente apresentados os 10 melhores hospitais, sempre para o total das admissões e algumas vezes com comparações entre casos cirúrgicos e médicos.

No Quadro XLVII são apresentados os 10 hospitais com valores mais elevados dos Índices de Casemix Composto, calculado pelos DRGs e calculado pelo Disease Staging para o total das admissões. Paralelamente são igualmente identificados os 10 hospitais com Índices de Casemix dos DRGs mais elevados em relação ao Disease Staging, os 10 hospitais na situação contrária e ainda os 10 hospitais com Índices de Casemix mais aproximados.

Quadro XLVII
10 Hospitais com Valores Mais Elevados dos
Índices de Casemix Composto, dos DRGs e do Disease Staging;
10 Hospitais com Maiores e Menores Diferenças entre os
Índices de Casemix dos DRGs e do Disease Staging
Total das Admissões

Composto	DRGs	Staging	> DRGs	> Staging	Semelhante
11	11	73	22	73	27
73	87	11	17	66	25
6	5	81	5	81	32
87	14	65	21	65	34
81	88	66	19	83	30
14	6	83	47	67	56
65	17	6	86	72	53
88	86	67	18	79	71
66	22	87	88	11	35
83	7	15	20	50	49

Assinalados a azul estão identificados os hospitais com Índices de Casemix mais elevados para as três formas de apuramento. Assinalados a vermelho estão identificados os hospitais com Índices de Casemix mais elevados, quando

calculados pelo Disease Staging e Composto. Assinalados a verde estão os hospitais com Índices de Casemix mais elevados quando calculados pelos DRGs e Composto.

Para os 3 hospitais com Índices de Casemix mais elevados, independentemente da forma de cálculo, 2 são do Tipo I e o restante do Tipo V. Para os hospitais comuns ao Disease Staging e ao Índice Composto 5 são do Tipo IV. Para os 2 hospitais comuns aos DRGs e ao Índice Composto, 1 é do Tipo I e o outro do Tipo V.

Para os 5 hospitais entre os 10 com Índice de Casemix mais elevado apurado pelos DRGs e que não aparecem nesta classificação no Índice de Casemix Composto, 2 são dos Tipos I e II e o restante do Tipo V. Quando se comparam os hospitais com Índice de Casemix mais elevado calculados pelo Disease Staging e não referidos no Índice de Casemix Composto, 1 é do Tipo I e o outro do Tipo IV.

Para os 10 hospitais com maiores diferenças entre os Índices de Casemix dos DRGs e do Disease Staging, com valores superiores para o primeiro, 6 são do Tipo II, 2 do Tipo V e 1 dos Tipos I e III. Na situação contrária, valores mais elevados com o Disease Staging, 8 hospitais são do Tipo IV e 1 dos Tipos I e III. Para os 10 hospitais com Índices de Casemix dos DRGs e do Disease Staging praticamente idênticos, 9 são do Tipo III e 1 do Tipo IV.

No Quadro XLVIII são apresentados os 10 hospitais com maior dispersão nas admissões (efeito raridade), em função da informação dos DRGs e do Disease Staging. Neste Quadro são igualmente identificados os hospitais com maiores e menores diferenças na dispersão da produção quando medidas pelos DRGs e pelo Disease Staging.

Quadro XLVIII
10 Hospitais com Maior Dispersão nas Admissões
(DRGs e Disease Staging)
10 Hospitais com Maiores e Menores Diferenças entre a Dispersão
Total das Admissões

DRGs	Staging	> DRGs	> Staging	Semelhante
9	10	42	---	11
10	13	23	---	73
13	9	59	---	19
2	1	62	---	20
8	25	87	---	18
61	2	45	---	17
25	26	57	---	22
26	3	58	---	51
1	61	82	---	21
3	8	49	---	83

Assinalados a azul estão os hospitais que têm maior dispersão na produção, tanto quanto esta é medida pelos DRGs, como pelo Disease Staging. Conforme se verifica para esta perspectiva os dois sistemas de classificação de doentes estão a dar a mesma informação, visto que são os mesmos hospitais. Destes 10 hospitais 7 são do Tipo I e 3 do Tipo III.

Para os 10 hospitais cuja concentração/dispersão da produção é maior nos DRGs em relação às indicações do Disease Staging, observa-se que nenhum pertence ao conjunto de hospitais com maior dispersão da produção (duas primeiras colunas do Quadro) e que 8 são hospitais do Tipo III, 1 do Tipo IV e 1 do Tipo V.

Conforme foi referido o Disease Staging apresenta níveis de concentração mais elevados quando comparados com os proporcionados pelos DRGs, pelo que não existe nenhum hospital em que a concentração/dispersão da produção tenha sido maior quando medida pelo Disease Staging.

Para os 10 hospitais em que o nível de concentração/dispersão da produção é praticamente idêntico para os dois sistemas de classificação de doentes, observa-se igualmente que nenhum é comum ao conjunto de hospitais com maior dispersão e que 6 hospitais são do Tipo II, 2 do Tipo IV, 1 do Tipo I e 1 do Tipo III.

No Quadro XLIX são apresentados os 10 hospitais com maior dispersão nas admissões, quando esta é medida pelo número de produtos diferentes responsáveis por 80% da produção. Mais uma vez é utilizada a informação dos DRGs e do Disease Staging. Neste Quadro são igualmente identificados os hospitais com maiores e menores diferenças na dispersão da produção quando medidas pelos DRGs e pelo Disease Staging.

Quadro XLIX
10 Hospitais com Maior Dispersão nas Admissões a 80%
(DRGs e Disease Staging)
10 Hospitais com Maiores e Menores diferenças entre a Dispersão
Total das Admissões

DRGs	Staging	> DRGs	> Staging	Semelhante
10	9	10	47	20
1	7	3	17	21
13	74	1	22	69
7	13	25		84
25	10	13		19
2	17	2		71
9	1	48		63
40	2	40		74
37	37	56		78
3	81	8		18

Assinalados a azul estão os hospitais que têm maior dispersão na produção, tanto quanto esta é medida pelos DRGs, como pelo Disease Staging. Existem 7 hospitais comuns aos dois sistemas de classificação de doentes, sendo 6 do Tipo I e 1 do Tipo III. Nos 3 hospitais somente identificados pelos DRGs 2 são do Tipo III e o restante do Tipo I, enquanto que para o Disease Staging 2 são do Tipo IV e 1 do Tipo II.

Para os 10 hospitais cuja concentração/dispersão da produção é maior nos DRGs em relação às indicações do Disease Staging, observa-se que 7 pertencem ao conjunto de hospitais com maior dispersão da produção quando esta é medida pelos DRGs e que 6 são hospitais do Tipo I e 4 do Tipo III.

Para os 3 hospitais cuja concentração/dispersão da produção é maior no Disease Staging em relação às indicações dos DRGs, observa-se que 1 (hospital 17) pertence ao conjunto de hospitais com maior dispersão da produção quando esta é medida pelo Disease Staging e que 2 são hospitais do Tipo II e 1 do Tipo III.

Para os 10 hospitais em que o nível de concentração/dispersão da produção é praticamente idêntico para os dois sistemas de classificação de doentes, observa-se igualmente que nenhum é comum ao conjunto de hospitais com maior dispersão e que 6 hospitais são do Tipo IV e 2 do Tipo II.

No Quadro L são apresentados os 10 hospitais com admissões mais complexas e graves, mais complexas e mais graves. Neste Quadro são igualmente identificados os hospitais com maiores e menores diferenças entre a complexidade e a gravidade dos casos admitidos.

Quadro L
10 Hospitais com Admissões Mais Complexas e Graves
10 Hospitais com Maiores e Menores Diferenças entre a
Complexidade e Gravidade dos Casos Admitidos
Total das Admissões

Total	Complexidade	Gravidade	> Complexidade	> Gravidade	Semelhante
6	13	73	9	20	63
14	6	14	10	18	51
87	9	11	13	63	76
15	87	81	88		19
5	10	83	87		81
88	88	65	26		49
7	15	70	6		68
86	14	77	1		28
13	86	80	86		78
73	7	66	7		45

Assinalados a azul estão identificados os hospitais com maior complexidade e gravidade dos casos, com maior complexidade e com maior gravidade, a vermelho os hospitais com presença simultânea no primeiro e terceiro atributos e a verde os hospitais com presença simultânea entre os dois primeiros atributos.

O hospital 14 (Tipo I) é o único que simultaneamente trata doentes mais complexos e mais graves, doentes mais complexos e doentes mais graves. Por sua vez o hospital 73 (Tipo IV) é igualmente o único que trata doentes mais complexos e mais graves e doentes mais graves.

Os hospitais 6, 7, 13, 15, 86, 87 e 88 (4 do Tipo I e 3 do Tipo V) são os que tratam simultaneamente doentes mais complexos e graves e doentes mais complexos.

O hospital 5 (Tipo I) embora figure entre os 10 que tratam doentes mais complexos e mais graves, não trata isoladamente doentes mais complexos ou doentes mais graves. O hospital 9 (Tipo I) somente aparece referenciado entre aqueles que tratam doentes mais complexos. Os hospitais 11, 65, 66, 70, 77, 80, 81 e 83 (7 do Tipo IV e o restante do Tipo I) somente aparecem referenciados entre as organizações de saúde que tratam doentes mais graves.

Nos 10 hospitais que comparativamente tratam mais doentes complexos do que graves, 6 são do Tipo I, 3 do Tipo V e 1 do Tipo III. Na situação contrária, em que a gravidade dos casos é superior à sua complexidade, somente são encontrados 3 hospitais, sendo 2 do Tipo II e 1 do Tipo IV. Finalmente, para os 10 hospitais cuja complexidade e gravidade dos casos tratados é mais semelhante são encontrados 5 hospitais do Tipo IV, 4 do Tipo III e 1 do Tipo II.

No Quadro LI são apresentados os 10 hospitais com maior efectividade para todas as admissões e para os casos cirúrgicos e médicos. Neste Quadro são igualmente identificados os hospitais com maiores e menores diferenças entre os casos médicos e cirúrgicos.

Assinalados a azul estão identificados os hospitais com maior efectividade para todas as admissões, para os casos cirúrgicos e para os casos médicos, a vermelho os hospitais com presença simultânea no primeiro e terceiro grupo de doentes e a verde os hospitais com presença simultânea entre os dois primeiros grupos de doentes.

Os hospitais 2, 9, 10, 41 e 61 (3 do Tipo I e 2 do Tipo III) aparecem entre os 10 hospitais com maior efectividade para todas as admissões e para os casos cirúrgicos e médicos, com a particularidade de o hospital 9 ocupar sempre a 1ª posição.

Os hospitais 2, 9, 10, 41 e 61 estão posicionados entre os 10 primeiros para o total das admissões e para os casos cirúrgicos e médicos.

Quadro LI
10 Hospitais com Maior Efectividade,
Todas as Admissões, Casos Cirúrgicos e Casos Médicos;
10 Hospitais com Maiores e Menores Diferenças entre os
Casos Cirúrgicos e os Casos Médicos

Todas as Admissões	Casos Cirúrgicos	Casos Médicos	> Cirúrgicos	> Médicos	Semelhante
9	9	9	88	13	45
13	48	13	87	1	18
61	61	10	48	44	14
10	87	61	62	40	42
2	86	2	59	60	33
15	41	15	25	7	24
48	39	44	12	36	31
32	2	7	55	16	56
41	10	32	39	34	69
86	25	41	29	15	19

Os hospitais 48 e 86 (1 do Tipo III e o outro do Tipo V) somente estão posicionados entre os 10 primeiros para o total das admissões e para os casos cirúrgicos, enquanto que os hospitais 13, 15 e 32 (2 do Tipo I e 1 do Tipo III) apresentam esta situação para o total das admissões e para os casos médicos.

Os hospitais 25, 39 e 87 (2 do Tipo III e 1 do Tipo V) somente estão posicionados entre os 10 primeiros para os casos cirúrgicos e os hospitais 7 e 44 (1 do Tipo I e o restante do Tipo III) para os casos médicos.

Para os 10 hospitais cuja efectividade nos casos cirúrgicos é bastante superior à dos casos médicos, 7 são do Tipo III, 2 do Tipo V e o restante do Tipo I. Na situação contrária, maior efectividade nos casos médicos, encontram-se 5 hospitais dos Tipos I e III. Com níveis de efectividade semelhantes para casos cirúrgicos e médicos encontram-se 6 hospitais do Tipo III, 2 do Tipo II e 1 dos Tipos I e 4.

No Quadro LII são apresentados os 10 hospitais com maior eficiência para todas as admissões e para os casos cirúrgicos e médicos. Neste Quadro são igualmente identificados os hospitais com maiores e menores diferenças entre os casos médicos e cirúrgicos.

Assinalados a azul estão identificados os hospitais com maior eficiência para todas as admissões, para os casos cirúrgicos e para os casos médicos, a vermelho os hospitais com presença simultânea no primeiro e terceiro grupo de doentes e a verde os hospitais com presença simultânea entre os dois primeiros grupos de doentes.

Quadro LII
10 Hospitais com Maior Eficiência,
Todas as Admissões, Casos Cirúrgicos e Casos Médicos;
10 Hospitais com Maiores e Menores Diferenças entre os
Casos Cirúrgicos e os Casos Médicos

Todas as Admissões	Casos Cirúrgicos	Casos Médicos	> Cirúrgicos	> Médicos	Semelhante
41	41	13	41	9	79
25	25	41	26	12	46
13	26	16	1	13	33
26	48	32	48	31	51
48	1	25	44	61	45
32	85	6	85	16	63
16	59	8	25	87	81
6	58	55	29	24	84
59	39	62	64	35	5
58	21	5	39	32	74

Os hospitais 25 e 41 (Tipo III) aparecem entre os 10 hospitais com maior eficiência para todas as admissões e para os casos cirúrgicos e médicos. Os hospitais 26, 48, 58 e 59 (Tipo III) recebem citações para todas as admissões e para os casos cirúrgicos, enquanto que para os hospitais 6, 13, 16 e 32 (3 do Tipo I e o restante do Tipo III) tal ocorre para todas as admissões e para os casos médicos.

Os hospitais 1, 21, 39 e 81 (1 dos Tipos I, II, IV e V) somente estão entre os 10 mais eficientes para os casos cirúrgicos e nos hospitais 5, 8, 55 e 62 (2 dos Tipos I e III) tal acontece nos casos médicos.

Para os 10 hospitais cuja eficiência é bastante superior nos casos cirúrgicos, 7 são do Tipo III, 2 do Tipo IV, sendo os restantes três distribuídos equitativamente pelos Tipos I e V. Para níveis de eficiência superiores nos casos médicos, encontram-se 5 hospitais do Tipo III, 4 do Tipo I e 1 do Tipo V. Entre os hospitais com níveis de eficiência similares para casos cirúrgicos e médicos estão 5 do Tipo IV, 4 do Tipo III e 1 do Tipo I.

No Quadro LIII são apresentados os 10 hospitais com melhor desempenho global para todas as admissões e para os casos cirúrgicos e médicos. Neste Quadro são igualmente identificados os hospitais com maiores e menores diferenças entre os casos médicos e cirúrgicos.

Assinalados a azul estão identificados os hospitais com melhor desempenho para todas as admissões, para os casos cirúrgicos e para os casos médicos, a vermelho os hospitais com presença simultânea no primeiro e terceiro grupo de doentes e a verde os hospitais com presença simultânea entre os dois primeiros grupos de doentes.

Quadro LIII
10 Hospitais com Melhor Desempenho Global,
Todas as Admissões, Casos Cirúrgicos e Casos Médicos;
10 Hospitais com Maiores e Menores Diferenças entre os
Casos Cirúrgicos e os Casos Médicos

Todas as Admissões	Casos Cirúrgicos	Casos Médicos	> Cirúrgicos	> Médicos	Semelhante
13	48	13	88	13	83
9	41	9	48	9	70
41	25	2	25	16	33
2	26	32	59	40	45
48	39	15	41	7	77
32	58	61	87	1	54
15	61	41	26	60	74
61	5	10	29	34	14
10	87	7	62	15	17
25	6	48	39	10	42

Os hospitais 41, 48 e 61 (Tipo III) estão sempre colocados entre os 10 melhores para qualquer dos agrupamentos de doentes considerados. Os hospitais 2, 9, 10, 13, 15 e 32 (5 do Tipo I e 1 do Tipo III) encontram-se entre os melhores para o total das admissões e para os casos médicos e no hospital 25 tal ocorre para o total das admissões e para os casos cirúrgicos.

Os hospitais 5, 6, 26, 39, 58 e 87 (3 do Tipo III, 2 do Tipo I e 1 do Tipo V) somente são referenciados para os casos cirúrgicos e o hospital 7 (Tipo I) para os casos médicos.

Para os 10 hospitais cujo desempenho global nos casos cirúrgicos é bastante superior ao dos casos médicos, 8 são do Tipo III e 2 do Tipo V. Quando o desempenho dos casos médicos é bastante superior ao dos cirúrgicos encontram-se 7 hospitais do Tipo I e 3 do Tipo V. Com desempenho global praticamente idêntico para os casos cirúrgicos e médicos encontram-se 4 hospitais dos Tipos III e IV e 1 dos Tipo I e II.

No Quadro LIV são apresentados para todas as admissões, para os casos cirúrgicos e para os casos médicos os hospitais com maiores e menores diferenças entre a efectividade e a eficiência.

Quadro LIV
Hospitais com Maiores e Menores Diferenças entre a
Efectividade e a Eficiência
Total das Admissões, Casos Cirúrgicos e Casos Médicos

Total			Cirúrgicos			Médicos		
> Efectividade	> Eficiência	Igual	> Efectividade	> Eficiência	Igual	> Efectividade	> Eficiência	Igual
9	12	69	9	1	76	9	12	5
61	59	78	61	26	68	10	59	18
10	62	63	87	16	84	44	62	68
2	55	77	86	40	29	61	55	11
44	25	5	10	13	20	13	88	17
13	30	11	4	41	51	2	30	22
15	1	82	48	59	53	36	16	32
4	16	64	88	33	42	15	25	45
36	33	45	22	12	81	60	87	58
86	26	3	2	25	70	54	33	63
								71
								74
								78
								82

Assinalados a azul estão os hospitais com comportamentos idênticos para o total das admissões e para os casos cirúrgicos e médicos, a verde encontram-se os hospitais nestas condições para os dois primeiros grupos e a vermelho para o total das admissões e para os casos médicos.

No que se refere ao grupo de hospitais com níveis de efectividade bastante superiores ao da eficiência podem destacar-se os seguintes aspectos:

- Os hospitais 2, 9, 10 e 61 (3 do Tipo I e 1 do Tipo III) apresentam sempre grandes disparidades entre a efectividade e a eficiência;
- Nos hospitais 13, 15, 36 e 44 (2 dos Tipos I e III) somente se observa este comportamento para o total das admissões e nos casos médicos, enquanto que nos hospitais 4 e 86 tal ocorre para o total das admissões e para os casos cirúrgicos;
- Nos hospitais 22, 48, 87 e 88 (2 dos Tipos V e 1 dos Tipos II e III) somente se verificam grandes diferenças nos casos cirúrgicos e nos casos médicos tal é observado nos hospitais 54 e 60 (Tipo III).

Para o grupo de hospitais com níveis de eficiência bastante superiores ao da efectividade observa-se o seguinte:

- Nos hospitais 12, 16, 25, 33 e 59 (2 do Tipo I e 3 do Tipo III) este comportamento ocorre em todas as situações;

- Nos hospitais 30, 55 e 62 (Tipo III) tal ocorre para todas as admissões e para os casos médicos e nos hospitais 1 e 26 (1 dos Tipos I e III) para o total das admissões e para os casos cirúrgicos;
- Nos hospitais 13, 40 e 34 (2 do Tipo III e 1 do Tipo I) este comportamento somente se verifica nos casos cirúrgicos e nos hospitais 87 e 88 (Tipo V) nos casos médicos.

Para o grupo de hospitais sem grandes discrepâncias entre a efectividade e a eficiência podem referir-se os seguintes aspectos:

- Não existe nenhum hospital em que tal se observe para os três grupos em estudo, nem para o total das admissões e para os casos cirúrgicos;
- Nos hospitais 5, 11, 45, 63, 78 e 82 (3 do Tipo IV, 2 do Tipo I e 1 do Tipo III) este comportamento é observado para o total das admissões e para os casos médicos;
- Os restantes hospitais somente aparecem referenciados em uma das perspectivas de análise, com a particularidade de nos hospitais 3, 64, 69 e 77 (3 do Tipo IV e 1 do Tipo I) tal ocorrer somente para o total das admissões e de que nenhum dos hospitais referenciados nos casos cirúrgicos é comum ao total das admissões e aos casos médicos.

Finalmente, será apresentado um resumo do desempenho por hospital incluindo os seguintes aspectos (ver Quadro LV):

- Identificação das situações, total das admissões, casos cirúrgicos, casos médicos e doenças (grandes agrupamentos de doenças) em que o hospital apresenta o melhor desempenho ou está incluído entre os cinco melhores. O resumo por desempenho respeita às três dimensões – global, efectividade e eficiência;
- Identificação das mesmas situações e das mesmas dimensões para os hospitais com pior desempenho, ou entre as cinco organizações de saúde com pior desempenho.

Como nota suplementar deve referir-se que para o total das admissões e para os casos cirúrgicos e médicos são mencionados os 10 melhores ou piores hospitais.

Atendendo igualmente a que todos os aspectos referentes ao desempenho global dos hospitais, bem como à respectiva efectividade e eficiência, foram apresentados e discutidos anteriormente, estes elementos são apresentados sem quaisquer comentários.

As Doenças estão agrupadas pelos Grandes Agrupamentos de Doenças do Disease Staging (ver Anexo I para visualizar as correspondência entre as siglas e as respectivas doenças).

Quadro LV
Síntese do Desempenho Hospitalar
Primeira Posição e Cinco Primeiras Posições
Última Posição e Cinco Últimas Posições
Total das Admissões, Casos Cirúrgicos e Médicos e Doenças

Hospital 1

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
OP	HB		HB	OP	Cir		Cir	Cir	CN	ND	BR
	IM		LY	VS	DE		CN	MS	OB		CV
	LY		OP		GI		GY		VS		MG
	PS		PN		HB		MS		OT		PN
			PS		IM		OB				
			RS		LY		OT				
					MS						
					PS						

Hospital 2

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
GY	Tot	GY	Tot	GY	DE		VS				HB
RS	Med		Cir	LY	HE						VS
	BR		Med		RS						
	EN		BR								
	IM		CV								
	OP		EN								
	PN		IM								
	RN		OP								
	SY		PN								
			RN								
			RS								
			SY								

Hospital 3

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
MG	BR	MG	BR		CV		OP		OP	PS	
	LY		PN		GI		PS				
	PN				HB						
					LY						
					ND						
					SY						
					OT						

Hospital 4

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
HE	DE	DE	IM	HE	RN		PN	VS	LY	CN	Tot
		HE		MG			VS		PN	OP	Cir
											Med
											GI
											IM
											PN

Hospital 5

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
RN	CV	RN	VS	RN	Med						
VS	LY				CV						
					VS						

Hospital 6

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Cir		DE	HB	Tot						ND
	HB		LY	RS	Med						
	MG		MG		HE						
	RS				PN						

Hospital 7

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Med		Med		IM	OP		OP			HE
	IM		IM		RN						ND
	PN		MS								
	PS		PN								

Hospital 8

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	MS		MS	CV	Med		EN			EN	VS
	OB		OB		IM						
					OP						

Hospital 9

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
CV	Tot	Tot	BR	NN	CV	RS	MS			Tot	Cir
EN	Med	Cir	GY		DE					HB	CN
GI	BR	Med	HE		HE					LY	GI
HB	DE	CV	PS		OP					MS	GY
IM	HE	EN	SY		SY					OB	MG
ND	SY	GI			VS					PN	PS
PN	VS	HB								RN	
NN		IM								RS	
		LY									
		ND									
		PN									
		VS									
		NN									

Hospital 10

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
OP	Tot	CN	Tot		CN	OB	IM	OB	IM	DE	Cir
	Med	OP	Cir		IM		PS		OT	GI	BR
	CN		Med				OT				EN
	CV		CV								HB
	LY		GI								MG
	ND		LY								OP
	VS		ND								PS
			VS								OT
			NN								

Hospital 11

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	IM										

Hospital 12

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
						Tot	EN	Tot	Cir		Tot
						Cir	HB	Med	DE		Cir
						Med	IM	CN	GI		DE
						CN	ND	CV	HB		EN
						CV	SY		MS		HB
						DE			ND		MS
						GI			SY		ND
						MS			VS		OP
						VS					SY

Hospital 13

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
Tot	EN	BR	Tot	Med	Tot		Cir		Cir		LY
Med	GI	PS	Med	CN	MS		DE		DE		
BR	HB	SY	CN		ND		MS		MS		
CN	MG		EN		SY						
PS	ND		GI		VS						
SY	OP		HB								
	VS		MG								
	OT		ND								
			OP								
			VS								
			OT								

Hospital 14

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							VS			CV	
										VS	

Hospital 15

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Tot		Tot	OT	RN		RS			IM	RS
	Med		Med								
	GI		GI								
	HE		HE								
	ND		IM								
	PS		ND								
	RN		OT								
	OT										

Hospital 16

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	HE		PS	BR	Tot		Cir		Cir		
				IM	Med		MG		CN		
					CN				EN		
					NN				MG		
									MS		

Hospital 17

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
											MS

Hospital 18

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	GY				GY						
					OT						

Hospital 19

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	GY				GY						DE
											EN
											RN
											RS

Hospital 20

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
OB	GY	OB			BR						OB
					GY						

Hospital 21

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
				EN	Cir						
				PS	MG						

Hospital 22

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	PN		MS		PN						

Hospital 23

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							HE		HE		
									RS		

Hospital 25

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
MS	Tot		Cir	MS	Tot		IM		CV		IM
	Cir		HE	ND	Cir				EN		
	RN		RN		Med				HB		
	RS		RS		BR				IM		
	SY		SY		CN						
					GI						
					RS						
					VS						

Hospital 26

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
DE	Cir		HE	DE	Tot	IM	OP	IM	GY		IM
	GI		PS	GI	Cir	NN		PN	MG		
	HB			OB	HB			NN	OB		
	HE			PN	MG				RS		
	ND				ND						
	PS				PS						
	VS										

Hospital 27

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							RN			MG	Tot
											Med
											CV
											OB
											RN
											VS
											NN
											OT

Hospital 28

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							PS		PS		

Hospital 29

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							Tot		Tot		Tot
							Med		Med		Med
							BR		GY		CN
							CN		PS		CV
							CV				GI
							LY				HE
							OB				LY
							PN				RN
							SY				SY
											VS

Hospital 30

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							Tot		Tot		
							Med		Med		
							CN		CN		
							CV		HE		
							ND		ND		
							OP		OP		
									OT		

Hospital 31

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					PS						Cir
											PN

Hospital 32

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Tot		Tot		Tot						
	Med		Med		Med						
	CN				ND						
	CV				RS						
					SY						

Hospital 33

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					BR	RN	CV	RN	Tot		
							ND		Cir		
							SY		Med		
									CV		
									ND		
									SY		

Hospital 34

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	MG		MG		OP				CV		
	OP		OP								
	RN		RN								

Hospital 35

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
LY	NN		GY		EN						
			LY		HE						
					LY						

Hospital 36

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							Tot	DE	Cir	Med	Tot
							Cir		MS		Cir
							DE				CV
							GY				HB
											MG
											ND
											VS

Hospital 37

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
									GY		

Hospital 38

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					CV		Cir	OT			Tot
							HE				Med
							OT				HE
											MS
											OP

Hospital 39

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Cir		Cir		Cir		OB		OB		
					OB						

Hospital 40

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
			OB			HB	Tot	HB	Tot		Cir
							Cir		Cir		Med
							Med		DE		IM
							DE		OP		MS
							MS		SY		
							OP		OT		
							PN				
							SY				
							OT				

Hospital 41

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Tot		Tot	Tot	Med						
	Cir		Cir	Cir	BR						
	Med		Med		DE						
	CV		OB		EN						
	GI				GI						
	MS				HB						
	OB				MG						
	RS				MS						
	OT				OB						
					OP						
					RS						

Hospital 43

Melhor Desempenho						Pior Desempenho					
Global		Efetividade		Eficiência		Global		Efetividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
			GY			PN	Tot		Tot	GY	Tot
							Cir		Cir		Med
							Med		Med		CN
							CN		CN		PN
							RS		PS		RS

Hospital 44

Melhor Desempenho						Pior Desempenho					
Global		Efetividade		Eficiência		Global		Efetividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Med						GI		GI	HE	Tot
	CN						RS				Med
											CN
											EN
											LY
											RS
											SY
											OT

Hospital 45

Melhor Desempenho						Pior Desempenho					
Global		Efetividade		Eficiência		Global		Efetividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							NN				NN

Hospital 46

Melhor Desempenho						Pior Desempenho					
Global		Efetividade		Eficiência		Global		Efetividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
			OB				GY		PN		GY
			RS								OB
			NN								

Hospital 48

Melhor Desempenho						Pior Desempenho					
Global		Efetividade		Eficiência		Global		Efetividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
Cir	Tot		Tot		Tot		IM		IM		
	EN		Cir		Cir				PN		
	GY		EN		EN						
	MG		HB		GY						
	MS		MG		MG						
			MS								

Hospital 51

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
								EN			

Hospital 52

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							Cir		Cir		

Hospital 54

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	OB						PN		PN		Med

Hospital 55

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	BR				Med		Tot	GI	Tot		LY
	OB				CV		Med		Med		
					PN		GI		HB		
					PS		HB		LY		
							LY		RN		
							RN		RS		
								SY			

Hospital 56

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					OB		MG		BR		

Hospital 57

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							MG				

Hospital 58

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Cir		NN		Tot						
	CN				Cir						
	NN				CN						
					EN						
					NN						

Hospital 59

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	MS				Tot		Tot		Tot		
					Cir		Med		Med		
					MS		CV		BR		
							GI		CV		
							HB		GI		
							HE		HB		
							ND		HE		
							PS		MG		
							RN		ND		
									PS		
									RN		
									RS		
									VS		
									NN		

Hospital 60

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
			CV				Cir		Cir		Tot
							EN		OP		Cir
							HB		VS		Med
							VS				DE
											GI
											PS
											NN
											OT

Hospital 61

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	Tot	MS	Tot				OB		OB		Cir
	Cir	RS	Cir								GY
	Med		Med								OB
	CN		CN								OP
	DE		DE								RN
	RS		GI								
			HB								
			RN								
			SY								
			OT								

Hospital 62

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					Med	ND	Tot	ND	Tot		
						PS	Med	PS	Med		
							BR	RS	BR		
							RS		GI		
									IM		
									RN		

Hospital 64

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
											PS

Hospital 65

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	DE		DE								

Hospital 66

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							NN		NN		

Hospital 67

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
									NN		

Hospital 68

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
											NN

Hospital 71

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					PN						

Hospital 75

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							Med				

Hospital 78

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							NN			NN	

Hospital 79

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							NN		NN		

Hospital 83

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
							DE		DE		

Hospital 84

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	NN		CN		NN						

Hospital 85

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
					Cir						
					OB						

Hospital 86

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
	EN		Tot		OT		BR		LY	BR	DE
	OT		Cir				LY				SY
			BR								
			DE								
			EN								
			GY								
			OT								

Hospital 87

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
OT	Cir	OT	Cir	SY	LY	MG	BR	MG	Med		BR
	SY		NN		NN		EN		BR		
	NN				OT		GY		EN		
							HE		GY		
							LY		HE		
									LY		

Hospital 88

Melhor Desempenho						Pior Desempenho					
Global		Efectividade		Eficiência		Global		Efectividade		Eficiência	
Prim	5 Prim	Prim	5 Prim	Prim	5 Prim	Ultima	5 Ult	Ultima	5 Ult	Ultima	5 Ult
			VS		RN	BR	Tot	BR	Tot	SY	BR
						EN	Med	EN	Med	OT	GY
						GY	GI	GY	MG		HE
						HE	MG	HE	RN		
						LY	RN	LY			
						SY		SY			
						OT					

DISCUSSÃO

- Discussão Metodológica
- Discussão dos Resultados
- Implicações da Avaliação da Produção e do Desempenho no Sector da Saúde

6. Discussão

O capítulo Discussão está organizado da seguinte forma:

- Discussão Metodológica
- Discussão dos Resultados
 - Produção
 - Medição da Produção
 - Perfil das Admissões
 - Desempenho
 - Efectividade
 - Eficiência
 - Global
- Implicações da Produção e do Desempenho Hospitalar no Sector da Saúde

Na Discussão Metodológica serão abordadas as questões mais gerais deste estudo, pelo que na Discussão dos Resultados, serão igualmente discutidas algumas questões específicas dos assuntos em análise.

6.1. Discussão Metodológica

Em função das finalidades e objectivos deste estudo foram analisadas duas grandes áreas dos hospitais, a produção e o desempenho. Tendo em atenção algumas limitações do sistema de informação hospitalar, por um lado e a importância do internamento na actividade hospitalar, por outro lado, optou-se por delimitar o âmbito do estudo a este sector.

Neste sentido, torna-se relevante identificar e discutir as principais questões metodológicas relacionadas com a produção e com o desempenho dos hospitais, nomeadamente as seguintes:

- Delimitação do sector de avaliação;
- Razão de escolha destas áreas: vantagens e inconvenientes;
- Razões para a escolha das dimensões que qualificam a produção e o desempenho hospitalar: vantagens e inconvenientes;
- Definição dos indicadores e dos procedimentos para medir as dimensões analisadas: vantagens e inconvenientes;
- Escolha dos instrumentos necessários para operacionalizar a medição da produção e do desempenho hospitalar: vantagens e inconvenientes;
- Definição das fontes de informação necessárias para a realização do estudo: vantagens e inconvenientes.

Em seguida, face às questões identificadas serão explicitadas algumas recomendações tendo em vista a melhoria da avaliação da produção e do desempenho hospitalar.

6.1.1. Sector de Avaliação

Conforme foi referido o estudo delimita o seu âmbito de aplicação somente a um sector da actividade hospitalar – o internamento.

Esta situação para além de obedecer a razões operacionais e pragmáticas relacionadas com a exequibilidade do estudo, é justificada, tanto pela importância específica do internamento no conjunto da actividade hospitalar, como por questões associadas com a disponibilidade da informação existente.

Em relação ao primeiro aspecto – relevância do internamento – pode referir-se que em 1990 o internamento justificava 73% do total das despesas hospitalares, enquanto que em 1999 este valor passa para 64% (Orçamento e Contas do SNS, 1990 e 1999). Quando se considera somente a produção, estes valores são de 71% e 62% do total da produção ajustada dos hospitais (Costa e Sena, 2002).

No que se refere ao segundo aspecto – disponibilidade da informação – deve referir-se que até ao presente momento, tanto nas consultas externas, como nas urgências, a informação é predominantemente quantitativa e agregada, pelo que não existe qualquer possibilidade de se ajustar pelo risco a produção destes sectores de actividade e consequentemente também não se pode avaliar o desempenho dos hospitais de acordo com esta metodologia.

No entanto, apesar da importância específica e do relevo do internamento na actividade hospitalar, os quais têm justificado a existência de metodologias próprias para a avaliação do desempenho hospitalar, constituindo bons exemplos o modelo “Best Hospitals” (Hill, Winfrey e Rudolph, 1997 e O’Muircheartaigh et al, 2002) e o modelo “Risk Adjusted Quality Outcome Measures” (DesHarnais et al, 1997 e 2000), deve-se no futuro introduzir alterações no sistema de informação hospitalar para se poder considerar o ambulatório como um sector igualmente relevante para o desempenho hospitalar.

Esta situação para além de estar contemplada em modelos como o “One Hundred Top Hospitals” (HCIA, 1999 e Griffith, Alexander e Jelinek, 2002), ou no “Hospital Ratings” (NHS, 2002), por exemplo, está igualmente retratada na realidade portuguesa, visto que entre 1990 e 1999 se observou uma redução da importância do peso relativo do internamento no total da actividade hospitalar, o que pode ser atribuído ao aumento dos hospitais de dia, médicos e cirúrgicos (cirurgia de ambulatório).

Assim, sugerem-se os seguintes aspectos, como determinantes para a introdução do sector ambulatório na avaliação da produção e do desempenho hospitalar:

- Introdução de um ou vários sistemas de classificação de doentes que permitam a identificação e comparação de produtos, incluindo a gravidade dos doentes. A escolha deste(s) sistema(s) de classificação de doentes deve ser criteriosa, visto que para além de responder às principais finalidades, nomeadamente a produção, qualidade, gestão interna e avaliação do desempenho, deve ainda ser compatível com a informação existente no internamento;
- Atendendo à crescente importância dos hospitais de dia na actividade hospitalar, devem ser implementadas metodologias que permitam a identificação dos doentes que podem ser tratados em ambulatório, sendo para tal necessário cruzar informação respeitante a produtos (doenças) e à gravidade dos doentes. Somente a título exemplificativo, refere-se a existência de uma metodologia que ao utilizar simultaneamente os DRGs e o Disease Staging, permite identificar a população alvo para tratamento em hospitais de dia (Taroni et al, 1991; Conklin et al, 1994; Taroni et al, 1997 e Louis, 2003).

6.1.2. Áreas de Avaliação

Em termos gerais para se avaliar a produção e o desempenho hospitalar devem ser utilizados elementos respeitantes à estrutura, ao processo e aos resultados.

Em que estrutura pode ser definida pelas características da oferta (organizações de saúde), processo por aquilo que é feito ao doente e resultados pela forma como o doente responde aos cuidados que lhe são prestados (QMAS, 1997).

Segundo os mesmos autores (QMAS, 1997) a estrutura inclui aspectos como a oferta tecnológica, a manutenção do equipamento, o cumprimento de normas de segurança, o acesso físico às instalações, a certificação, treino e educação contínua do pessoal, o horário de operação, os esquemas de marcação de contactos e os sistemas de informação e de comunicação; no processo estão consideradas dimensões como a qualidade clínica ou técnica, a qualidade nas relações entre prestadores e doentes, o acesso aos cuidados de saúde e a adequação dos cuidados prestados e nos resultados estão incluídas a mortalidade, a doença, a incapacidade, o desconforto e a insatisfação.

Existe uma enorme controvérsia sobre quais as dimensões que devem ser utilizadas para se avaliar o desempenho hospitalar (ver revisão de literatura), no entanto deve ter-se presente duas afirmações de Donabedian – “A presença de uma associação válida entre o processo e os resultados significa que somente se podem obter determinados resultados em determinadas condições específicas. Não se pode inferir que os resultados observados são para uma determinada situação decorrentes de procedimentos anteriores” (Donabedian, 1980) e “Uma boa estrutura aumenta a probabilidade de se obter um bom processo e um bom processo aumenta a probabilidade de se obter um bom resultado” (Donabedian, 1986).

Por outro lado, deve ainda ter-se presente a definição de qualidade apresentada por Lohr (1990) “A qualidade dos cuidados exprime o nível dos serviços de saúde prestados a indivíduos ou populações no sentido de aumentar a probabilidade de ocorrer o resultado desejável, consistente com os conhecimentos técnicos”.

Tendo em atenção estas duas últimas afirmações, neste estudo optou-se por avaliar o desempenho através dos resultados de saúde em função da seguinte equação (QMAS, 1997):

$$\text{Resultados} = \text{Efectividade dos tratamentos existentes} + \text{Características dos Doentes} + \text{Qualidade dos Cuidados Prestados} + \text{Componente Aleatória}$$

As principais vantagens para utilizar a abordagem pelos resultados podem ser retirados dos artigos de lezzoni e outros (1992a e 1992e) e de lezzoni (1995), referidos na revisão de literatura:

- Acessibilidade da informação referente a resultados de saúde (mortalidade e demora média, por exemplo);
- Validade das medidas de resultados, mortalidade e demora média, como indicadores da efectividade e da eficiência dos cuidados prestados (ver revisão da literatura);
- Clareza na sua interpretação para todos os agentes intervenientes no mercado de saúde. Por exemplo, a mortalidade é facilmente definida e medida e constitui, na grande maioria das situações, um resultado indesejável para os doentes.

A estes argumentos pode ainda ser acrescentado outro respeitante à identificação genérica de eventuais problemas em determinado hospital, tornando-se assim mais dirigida a investigação na estrutura e no processo de eventuais problemas existentes nestas organizações de saúde (Walley, 1999).

Como principais argumentos contrários à utilização de indicadores de resultados para avaliar o desempenho dos hospitais, podem recordar-se as questões enunciadas por Thomas e Hofer (1998 e 1999), essencialmente quando afirmam que a taxa de mortalidade mesmo que ajustada pelo risco não é um bom indicador de qualidade, pelo que publicitação de resultados deste tipo pode contribuir para a desinformação do público sobre o desempenho dos hospitais e defendem que os indicadores de processo são mais adequados para se medir a efectividade dos cuidados prestados.

Embora com uma opinião mais reservada Lohr (1988) afirma que quando os resultados não estão associados à prática médica oferecem poucas indicações para a melhoria da qualidade dos cuidados prestados, pelo que disponibilizam argumentos contra a sua utilização para efeitos da avaliação e garantia da qualidade.

No citado estudo da QMAS (1997) é igualmente referido que existem situações em que relacionar os resultados com a qualidade dos cuidados prestados pode ser problemático, nomeadamente quando os resultados não podem ser relacionados com as intervenções, quando o doente contacta o hospital com “atraso” (a este respeito dão o exemplo do acidente cerebrovascular) ou ainda quando a “compliance” do doente não é conseguida.

Face ao exposto, deve ter-se presente que embora a abordagem pelos resultados seja na generalidade válida, pode conter alguns problemas que poderão contribuir para enviesar a apreciação do desempenho do hospital.

Assim, deve concluir-se que a abordagem pelos resultados é condição necessária, mas não suficiente para avaliar o desempenho hospitalar, embora deva constituir o elemento de partida para se concretizar aquele objectivo, pelo que se sugere que numa segunda fase se desenvolvam metodologias, modelos e indicadores para avaliar a estrutura e o processo.

Para atingir este objectivo sugere-se igualmente a consideração das metodologias preconizadas pelo “Best Hospitals” (EUA) e do “Hospital Ratings” do Reino Unido.

A avaliação de elementos da estrutura e do processo pode também conter dados relevantes para os doentes que, para além de reduzirem as assimetrias de informação existentes, fundamenta melhor o processo de escolha destes agentes.

Para caracterizar a actividade dos hospitais, embora se preconize a avaliação do desempenho utilizando os resultados, deve igualmente ser utilizada outra área – a produção. Em termos genéricos o “output” de uma determinada organização pode ser considerado como um resultado intermédio.

Contudo, atendendo a que na teoria sobre a avaliação dos resultados dos hospitais (ver revisão de literatura) é considerada uma divisão entre resultados finais e intermédios, sendo a mortalidade considerada um indicador de resultados finais e as complicações e readmissões como indicadores de resultados intermédios.

Neste sentido, duas interrogações se colocam:

- Como qualificar a produção do hospital?
- Quais as razões justificativas para a consideração da produção hospitalar?

Em relação ao primeiro aspecto propõe-se a introdução de mais uma vertente à trilogia definida por Donabedian (1985), sugerindo-se o seguinte esquema:

Estrutura – Processo – Produção (“Output”) – Resultados (“Outcome”)

Nesta conformidade, as principais razões justificativas para a consideração da produção são de dupla natureza:

- O valor intrínseco das medidas da produção hospitalar para o ajustamento pelo risco e para a avaliação do seu desempenho (plano conceptual);
- A relação entre a produção e o desempenho hospitalar (plano operacional).

Como foi referido no Capítulo Enquadramento Teórico e Questões Práticas o produto do hospital respeita à prestação de cuidados de saúde em função de um determinado resultado (Hornbrook, 1982), pelo que o episódio de internamento configura a melhor medida para se definir o produto hospitalar.

Nesta conformidade, atendendo igualmente às abordagens possíveis para se definirem, identificarem e compararem casos, os sistemas de classificação de doentes ou os índices escalares (também designados por Índices de Casemix), a consideração da primeira alternativa configura-se como o principal “input” para a avaliação do desempenho hospitalar.

Isto porque, para se avaliar o desempenho é necessário ajustar pelo risco, ou seja identificar as características dos doentes relevantes para conduzirem a diferentes resultados de saúde.

Atendendo à hierarquia proposta por Hornbrook (1982) para se estabelecerem critérios hierarquizados para definir produtos (ver Enquadramento Teórico e Questões Práticas) facilmente se compreende que no topo desta hierarquia se encontra a definição de produtos com resultados semelhantes.

Por outro lado, como foi igualmente referido, os dois aspectos mais relevantes para o ajustamento pelo risco respeita à forma como se integram as diferentes dimensões do risco, o que na prática tem conduzido à criação de diversos sistemas de classificação de doentes que pretendem identificar e medir a gravidade dos doentes.

Assim, torna-se lógico concluir que a definição e a medição da produção constituem, em si mesmas, um elemento de partida para definir e operacionalizar modelos de avaliação do desempenho hospitalar.

Neste sentido, torna-se igualmente importante discutir a validade, bem como as vantagens e inconvenientes dos sistemas de classificação de doentes utilizados, essencialmente no que respeita às suas propriedades para medir a produção hospitalar e avaliar o desempenho hospitalar. Esta temática será abordada num ponto posterior da Discussão Metodológica, quando se analisarem os instrumentos do estudo.

A segunda questão, de natureza mais operacional, pretende suscitar a discussão sobre os eventuais impactes de diferentes níveis produtivos, tanto em termos quantitativos, como qualitativos, no desempenho dos hospitais.

As questões quantitativas são debatidas na literatura internacional, tanto para a efectividade dos cuidados prestados (Blumberg, 1986; Farley e Ozminkowski, 1992; Iezzoni et al, 1997; Bach et al, 2001; Birkmeyer et al, 2002), como para a eficiência (Wagstaff, 1989; Paiva, 1993; Arndt, Bradbury e Golec, 1995; Butler, 1995; Given, 1996; Wholey et, 1996; Dranove, 1998; Barros e Sena, 1999; Cabral e Barriga, 1999 e Carreira, 1999).

No caso da efectividade as questões mais relevantes respeitam à existência ou não de associação entre a quantidade produzida e optimização de resultados e na eficiência, a literatura com um conteúdo mais económico, pesquisa a existência de economias de escala e concomitantemente de uma dimensão óptima de produção.

Para as questões qualitativas, os aspectos mais relevantes respeitam aos eventuais impactes da selecção adversa ou de escolha de doentes mais favoráveis (normalmente menos complexos e menos graves) sobre os resultados dos hospitais.

Estes aspectos são identificados no Capítulo Enquadramento Teórico e Questões Práticas, pelo que neste momento somente se refere a importância de diferentes níveis produtivos dos hospitais sobre o seu desempenho.

Assim, pode concluir-se que a consideração da produção quando se analisa o desempenho hospitalar é perfeitamente justificada, visto que existe uma associação forte entre estas duas áreas, tanto na alimentação dos modelos de desempenho, como na justificação de eventuais diferenças nos resultados dos hospitais.

Finalmente, deve ainda referir-se que embora a produção constitua, por si mesma, uma área importante, visto que pode exprimir o esforço dos hospitais para corresponder às necessidades dos indivíduos e das populações, é relevante identificar as consequências desta mesma produção e ainda a adequação da sua prestação.

Enquanto, que o primeiro aspecto é apreciado com a avaliação do desempenho dos hospitais, o segundo necessita de metodologias específicas para a sua análise.

Atendendo às exigências de um estudo deste tipo, estes aspectos não serão aqui analisados. No entanto, atendendo às características do Serviço Nacional de Saúde português e às recentes reformas introduzidas no sector hospitalar, sugere-se que no futuro se realizem estudos específicos sobre a adequação dos internamentos, o que para além de incluir os aspectos referidos para a identificação de situações que podem ser tratadas em ambulatório, exige ainda a definição de critérios explícitos, por doença, para a justificação de internamento ou identificação de métodos alternativos de tratamento.

Sem pretender esgotar o assunto, deve no entanto evidenciar-se que, por um lado, a consideração da gravidade dos doentes é um aspecto essencial e que, por outro lado, é necessária a constituição de equipas de peritos para criar os critérios para a adequação dos cuidados prestados.

6.1.3. Dimensões

Definido o sector em estudo – o internamento – bem como as áreas que foram analisadas – a produção e o desempenho hospitalar – é agora importante discutir as questões associadas com as dimensões em apreciação.

No estudo foram utilizadas as seguintes dimensões:

Figura 54
Dimensões

Produção	Desempenho
<ul style="list-style-type: none">• Medição da Produção• Perfil das Admissões	<ul style="list-style-type: none">• Efectividade• Eficiência• Global

Para a medição da produção foram consideradas duas dimensões, a medição da produção e o perfil das admissões.

A concretização da medição da produção hospitalar exige a combinação de dois vectores: a definição de um sistema de identificação de produtos ou de classificação de doentes, como é frequentemente designado na literatura internacional e a definição de um sistema (metodologia) que permita ponderar e comparar os produtos (Costa, 1994).

Conforme foi referido, com os sistemas de classificação de doentes pretende-se identificar o “casemix” dos hospitais, ou seja a variedade das situações clínicas tratadas por cada hospital, organização de saúde ou prestador (Lichtig, 1986) e o Índice de Casemix que é o valor que quantifica a variedade dos casos tratados em cada hospital (Lichtig, 1986).

Neste sentido, embora existam diversas críticas à utilização de indicadores de síntese (ver Enquadramento Teórico e Questões Práticas e Áreas de Avaliação da Discussão Metodológica), torna-se importante medir a produção dos hospitais.

Em primeiro lugar, visto que enquanto o “casemix” traduz as diferenças reais nos produtos hospitalares, o Índice de Casemix é um indicador que permite ponderar os episódios de internamento tratados nestas organizações de saúde e igualmente avaliar os seus resultados em função destas mesmas características dos doentes.

Neste estudo foram utilizados dois sistemas de classificação de doentes, os DRGs e o Disease Staging (ver Metodologia – Instrumentos e o ponto posterior da Discussão Metodológica – Instrumentos – para identificar, tanto as respectivas características, como as principais vantagens e inconvenientes).

Como os DRGs e o Disease Staging, para além de disponibilizarem informações sobre a concentração da produção hospitalar, estão ainda a medir dimensões diferentes dessa produção – a complexidade e a gravidade dos casos tratados – foi realizada uma análise comparativa da produção dos hospitais em função dos resultados destes sistemas de classificação.

Como é natural os Índices de Casemix da complexidade e da gravidade dos casos tratados por hospital estão a disponibilizar indicações distintas sobre a produção hospitalar, pelo que se sugeriu igualmente a utilização de um Índice de Casemix conjunto, calculado através da média não ponderada dos Índices de complexidade e de gravidade.

No entanto, atendendo à importância específica do assunto, bem como nas implicações que um sistema de medição da produção hospitalar tem sobre outras áreas, nomeadamente o financiamento e a avaliação da efectividade e da eficiência, sugeriu-se a necessidade de se realizarem estudos futuros, não só para avaliar as melhores metodologias para se medir complexidade e gravidade, mas também sobre a eventualidade de se introduzirem ponderações distintas para estas duas dimensões.

Para a avaliação do desempenho foram consideradas três dimensões, a efectividade dos cuidados prestados, a eficiência dos hospitais e uma medida global, ponderando com a mesma intensidade a efectividade e a eficiência.

Conforme foi referido a efectividade pretende medir os resultados ou consequências de determinado procedimento ou tecnologia médica quando aplicados na prática (McGuire, Henderson e Mooney, 1988; Pereira, 1993).

Trata-se de uma dimensão que assume particular importância para todos os agentes do mercado da saúde, com especial incidência para os consumidores e prestadores. Para os consumidores, dado que o conhecimento dos hospitais ou dos prestadores que prestam serviços com maior efectividade e qualidade, para além de poder reduzir as assimetrias da informação existentes, permite igualmente uma melhor fundamentação da escolha dos prestadores. Para os prestadores, nomeadamente os médicos, visto que a optimização dos resultados das suas intervenções constitui uma das suas principais preocupações.

Por outro lado, a eficiência, como foi igualmente referido, é uma medida que exprime a relação entre os recursos utilizados e os resultados obtidos (Pereira, 1993).

Trata-se de uma dimensão que tem preocupado os proprietários e os gestores, sendo particularmente ilustrativo deste comportamento, o actual debate existente internacionalmente e em Portugal (sobretudo em função dos hospitais SA). Contudo, tanto os médicos, como os consumidores, embora com menor intensidade, também consideram a eficiência como uma dimensão importante.

Para os consumidores, as principais preocupações assumem uma natureza dupla: em primeiro lugar, atentas as características do Serviço Nacional de Saúde português, porque pretendem reduzir as suas contribuições (impostos) para o Estado e em segundo lugar, quando se perspectiva a eficiência em termos de demora média, porque a redução na duração de internamento, desde que os resultados tenham qualidade, constituem uma das suas preferências (Murphy e Cluff, 1990; Iezzoni et al, 1996c e Phillips et al, 2000).

Finalmente, a consideração de uma medida conjunta do desempenho global permite a correspondência com as expectativas e preferências dos agentes do mercado da saúde e ainda a identificação dos melhores e piores hospitais, bem como a investigação sobre eventuais contradições para a optimização da efectividade e da eficiência.

Para que tal se concretize torna-se importante discutir qual a melhor metodologia para se avaliar a efectividade, a eficiência e o desempenho global dos hospitais.

Para além das áreas de intervenção sobre a qual deve incidir a medição da efectividade e da eficiência (ver Enquadramento Teórico e Questões Práticas e o discutido nas Áreas de Intervenção deste capítulo) é importante discutir a melhor metodologia para a sua concretização.

Durante este estudo foram identificadas e debatidas as questões teóricas (ver Capítulo Enquadramento Teórico e Questões Práticas) que sustentam a não consideração dos resultados brutos (mortalidade e demora média, por exemplo) como bons indicadores do desempenho dos hospitais, visto que os mesmos podem tratar um “mix” de doentes completamente distinto.

Foi também evidenciada durante a análise empírica (ver Capítulo Resultados) as discrepâncias no desempenho e na respectiva ordenação dos hospitais em função de resultados brutos e de resultados ajustados pelo risco.

Ajustamento pelo risco corresponde à realização dos procedimentos necessários para identificar as características dos doentes que podem conduzir a resultados diferentes (Iezzoni, 1989, 1996b e 1997c e Gross et al, 2000).

Também é internacionalmente pacífico que as metodologias para ajustamento pelo risco não devem conter elementos respeitantes às características dos hospitais, designadamente porque as opções de tratamento dos hospitais são o que deve ser avaliado e não servir de base à avaliação (Hornbrook, 1982 e 1985; Blumberg, 1986; Hughes et al, 1996; Thomas, Ashcraft e Zimmerman, 1986; Iezzoni et al, 1996; Iezzoni, 1997a; Iezzoni, 1997b; Iezzoni, 1997c; Desharnais et al, 1997 e 2000; Thomas e Hofer, 1998 e 1999).

Neste sentido, atendendo a que os DRGs, por um lado, são uma medida da complexidade dos casos tratados, sendo portanto uma medida mais posicionada

do lado da oferta do que da procura e a que, por outro lado, incluem critérios de utilização e exprimem o que foi feito e não o que deveria ter sido feito (Hornbrook, 1982 e 1985; Aronow, 1988; Costa, 1991 e 1994), não foram considerados para o ajustamento pelo risco.

A este propósito, como existem versões dos DRGs, designadamente os All Patient Refined e International Refined, será igualmente interessante considerar a sua utilização para se proceder ao ajustamento pelo risco. Voltar-se-á mais tarde a esta discussão, no momento em que se analisarão os Instrumentos do estudo.

Sendo assim, a efectividade e a eficiência, embora com metodologias distintas, foram avaliadas através da comparação entre os valores observados e esperados.

Para estimar os valores esperados utilizou-se a gravidade dos doentes, medida pelo Disease Staging, bem como a doença principal, a idade e as comorbilidades (também com estadios de severidade) dos doentes e ainda a forma de admissão.

O desempenho global, visto que corresponde à média simples entre a efectividade e a eficiência, foi calculado directamente dos valores destas duas dimensões, bastando para tal assegurar que a efectividade e a eficiência estejam na mesma escala (voltar-se-á a este assunto na Discussão dos Indicadores).

6.1.4. Indicadores e Procedimentos

Definido o sector em estudo – o Internamento –, bem como as áreas que foram analisadas – a produção e o desempenho hospitalar – e ainda as dimensões em estudo – Medição da Produção, Perfil das Admissões, Efectividade, Eficiência e Global / Efectividade e Eficiência) é agora importante discutir os indicadores utilizados no estudo.

No estudo foram utilizados os seguintes indicadores:

Figura 55
Indicadores

Produção		Efectividade	Desempenho	
Medição da Produção	Perfil das Admissões		Eficiência	Desempenho Global
Índice de Casemix pelos DRGs	Concentração Total	Relação entre a mortalidade observada e a mortalidade esperada (mortalidade ajustada pelo risco)	Relação entre a demora média observada e a demora média esperada (demora média ajustada pelo risco)	Média simples da mortalidade ajustada pelo risco e da demora média ajustada pelo risco
Índice de Casemix pelo Disease Staging	Concentração a 80%			
Índice de Casemix conjunto	Complexidade			
	* Níveis			
	* Global não Ponderada			
	* Global Ponderada			
	Gravidade			
	* Níveis			
	* Global não Ponderada			
	* Global Ponderada			
	Conjunto (Complexidade e Gravidade)			
	Diferença entre a complexidade e a gravidade			

O Índice de Casemix foi utilizado como indicador de síntese da actividade hospitalar (ver Capítulo Enquadramento Teórico e Questões Práticas) tendo em atenção a informação disponibilizada pelos dois sistemas de classificação de doentes utilizados no estudo.

Assim, com os DRGs está-se a medir a complexidade dos casos tratados em cada hospital, enquanto que o Disease Staging traduz a gravidade dos casos tratados em cada hospital.

Em relação aos DRGs, atentas as características específicas deste sistema de classificação, a utilização de um comparador para os diferentes produtos (peso

específico) tem de ser feita de forma mediata, visto que os DRGs não contêm nenhum elemento que permita uma utilização directa (Costa, 1994).

Assim, este elemento comparador pode ser calculado de duas formas (Hornbrook, 1982 e Lichtig, 1986): em função da demora média observada de cada DRG; ou através dos pesos específicos dos DRGs quando utilizados para financiamento.

A utilização da demora média observada tem os seguintes problemas (Costa, 1994):

- Reflecte o que foi feito e não o que deveria ter sido feito;
- Não considera a intensidade da utilização de recursos, pelo que DRGs com demoras médias idênticas têm sempre o mesmo peso específico;

Quando se utiliza o peso específico constante nas tabelas de financiamento podem ser referidos os seguintes problemas (Costa, 1994):

- Reflecte o que foi feito e não o que deveria ter sido feito;
- Considera que todos os produtos (DRGs) têm a mesma eficiência na utilização de recursos, visto que os preços devem ter correspondência com os custos médios;
- Pressupõe que os preços traduzem, por uma lado, as preferências dos consumidores e, por outro lado, que existe uma relação entre preços e qualidade dos cuidados prestados.

Neste estudo opta-se pela segunda alternativa – utilização dos pesos específicos dos DRGs – para se compararem os diversos produtos.

Em primeiro lugar, porque o Índice de Casemix actualmente utilizado segue esta metodologia (Bentes et al, 1996), pelo que se torna um indicador mais familiar na realidade portuguesa.

A este propósito, como foi referido, para minimizar o efeito de neutralidade económica, utiliza-se a fórmula bruta para cálculo do Índice de Casemix, a qual não dá ponderações distintas para internamentos de curta duração, de longa duração e de evolução prolongada.

Em segundo lugar, porque a utilização da demora média, atentos os diferentes custos entre os produtos, pode originar informações enviesadas, embora pareça não existir grande evidência empírica sobre o assunto (ver Capítulo Enquadramento Teórico e Questões Práticas).

Para o Disease Staging, embora a ponderação da produção seja igualmente obtida de forma mediata quando se considera a versão original, a utilização do

“software” do Disease Staging e essencialmente as suas escalas de previsão de mortalidade (as quais foram recalibradas a Portugal) pode referir-se que o elemento comparador (peso específico) é calculado de forma directa.

No estudo utiliza-se igualmente um Índice de Casemix Conjunto (com ponderação idêntica para os dois índices anteriormente referidos – calculados através dos DRGs e do Disease Staging), o qual para além de pretender sintetizar a complexidade e a gravidade dos casos tratados, pretende igualmente valorizar questões associadas com a oferta e com a procura e ainda com o que foi feito e o que deveria ser feito.

Finalmente, deve referir-se que se subscrevem todas as dúvidas e reservas apontadas aos índices escalares, essencialmente a perda de informação e de sensibilidade quando se utiliza uma medida única para sintetizar a produção hospitalar (Hornbrook, 1992).

No entanto, estas mesmas razões, utilização de uma medida única, são igualmente justificativas da sua consideração para a medição da produção hospitalar.

Tal não deve ser entendido que, por um lado, não devem ser utilizados outros indicadores para se caracterizar a produção e que, por outro lado, não se realizem investigações futuras e mais específicas sobre as dimensões que devem qualificar o “casemix” dos hospitais, bem como sobre a respectiva ponderação.

A concentração da produção, como foi referido, envolve duas perspectivas: a concentração de produtos entre hospitais (concentração vertical) e a concentração de determinados produtos num número pequeno de hospitais (concentração horizontal) (Evans e Walker, 1972 e Schumaker, 1979).

Neste estudo, atendendo aos seus objectivos, a concentração horizontal não é analisada, considerando-se somente os aspectos associados à concentração vertical.

Para tal foram considerados dois indicadores: o número de produtos diferentes em cada hospital em função do número de produtos diferentes em todos os hospitais e o número de produtos responsável por 80% da produção.

A principal razão para se analisar a concentração da produção nos hospitais é encontrada no argumento se a concentração, ou o seu contrário a raridade dos casos, é ou não por vezes confundida com a complexidade ou com a gravidade dos casos tratados (Tatchell, 1983).

O estudo prova que a raridade dos casos (contrário da concentração) pode por vezes estar a ser confundida com a gravidade dos casos. Isto porque, os

hospitais de maior dimensão, aqueles que estão no fim da linha do sistema de referência, são os que apresentam menor concentração (mais casos raros), maior complexidade dos casos, mas não acontecendo o mesmo em relação à gravidade.

O âmbito destes resultados, bem como a pesquisa de eventuais associações entre concentração e resultados em saúde (os quais não foram abordados no estudo) merecem igualmente que no futuro se realizem estudos sobre o assunto.

A complexidade dos casos, como foi referido, é um indicador que deve sintetizar a quantidade de recursos necessários para tratar cada doente (Luke, 1979), pelo que está mais associado com as características dos hospitais, essencialmente o desenvolvimento tecnológico e com a especialização dos recursos humanos (Hornbrook, 1992). A este propósito também foi referido que os DRGs, atentas as suas especificidades, são uma boa medida da complexidade dos casos tratados (McGuire, 1991).

A consideração deste indicador tem uma dupla utilidade: em primeiro lugar permite identificar os hospitais que tratam doentes mais complexos e, em segundo lugar, permite associar a complexidade os casos tratados com os resultados dos hospitais.

O primeiro aspecto – identificação de hospitais – pode disponibilizar informações sobre a distribuição de casos por hospital, com as naturais consequências ao nível do planeamento dos recursos tecnológicos e humanos, como na gestão interna dos hospitais (indicador de oferta).

Com o segundo aspecto – associação com os resultados – pretende-se identificar hospitais cujos resultados (bons ou maus) possam ser atribuídos (descontados) às características da sua procura.

Neste estudo a complexidade dos casos foi medida de três formas: através de níveis de complexidade, através de um indicador global e não ponderado de complexidade e de um indicador global e ponderado de complexidade.

Para o primeiro indicador – níveis de complexidade – foram utilizados os pesos relativos dos DRGs. Para o segundo indicador – medida global não ponderada – foi construído um “z score” em função dos níveis de complexidade definidos. Para o terceiro indicador – medida global ponderada – foi utilizado o Índice de Casemix dos DRGs.

Os dois primeiros indicadores pretendem somente descrever o perfil das admissões dos hospitais, enquanto que o terceiro pretende analisar implicações deste mesmo perfil sobre o desempenho dos hospitais.

A necessidade de se utilizarem, neste estudo, estes três indicadores foi a seguinte:

- Em primeiro lugar, é necessário descrever e compreender o perfil das admissões;
- Em segundo lugar, é necessário comparar esse mesmo perfil das admissões entre os diversos hospitais;
- Em terceiro lugar, é necessário que esta comparação seja feita em função das diferentes doenças, as quais podem ter diferentes intensidades, apesar de estarem situadas nos mesmos níveis de complexidade.

Neste sentido, conforme foi referido, para os dois primeiros aspectos, tanto a utilização dos níveis de complexidade, como o “z score” parecem adequados. Para o terceiro aspecto, o “z score” poderia ser utilizado se todas as doenças tivessem a mesma complexidade dentro do mesmo nível.

Como tal comportamento não é teoricamente esperado, ensaiou-se a utilização do Índice de Casemix dos DRGs como “proxy” da complexidade. Atendendo, a que este Índice de Casemix e o “z score” estão significativamente correlacionados, optou-se pela utilização apenas do Índice de Casemix para análises relacionais.

A gravidade dos casos permite identificar as características dos doentes que podem conduzir a resultados menos favoráveis. Embora exista uma enorme variedade de definições para a gravidade dos casos, neste estudo utiliza-se a gravidade como probabilidade de morte ou de falência de um órgão (Thomas, Ashcraft e Zimmerman, 1986).

À semelhança do que foi referido para a complexidade a utilidade do indicador é igualmente dupla, com identificação dos hospitais que tratam doentes mais graves e a associação desta gravidade com os resultados dos hospitais.

Por outro lado, mais uma vez à semelhança do que foi efectuado para a complexidade foram utilizados três indicadores para medir a gravidade.

Atendendo a que, os problemas são conceptualmente similares aos referidos para a complexidade, praticamente a única diferença existente é entre a utilização da gravidade em vez de complexidade, as razões da sua utilização não são aqui discutidas.

Se para esta última questão são mantidas as observações efectuadas para a complexidade, a identificação dos hospitais com doentes mais ou menos graves assume contornos diferentes, visto que se trata predominantemente de um indicador de procura.

Assim, é interessante investigar se os hospitais que tratam doentes mais complexos, também tratam doentes mais graves e na eventualidade de tal não acontecer descortinar as principais razões para tal comportamento.

Na realidade portuguesa, esta questão parece ainda mais relevante, visto que a universalidade dos cuidados, bem como a inexistência de barreiras à acessibilidade (pelo menos para os doentes da área de atracção) parece traduzir-se por uma associação directa entre a complexidade e a gravidade dos casos.

Este é igualmente o discurso oficial e informal, no qual se afirma que os hospitais universitários e centrais, por estarem no fim da linha da referência dos doentes, são os que tratam doentes mais complexos e mais graves.

Os indicadores – concentração, complexidade e gravidade – permitem caracterizar o perfil das admissões hospitalares e inclusivamente a política das admissões hospitalares, nomeadamente a existência ou não de práticas de selecção adversa.

Actualmente em Portugal, esta questão pode ainda ganhar mais relevo, visto que frequentemente se refere que as novas experiências de gestão hospitalar contêm incentivos para que os hospitais escolham os doentes.

Tal comportamento deverá no futuro ser analisado, avançando-se desde já com mais uma recomendação – a pesquisa de eventuais políticas de admissões distintas nos hospitais em determinado momento deverá ser sempre comparada com os comportamentos existentes anteriormente.

Conforme se referiu, neste estudo as práticas de selecção adversa não serão analisadas, identificando-se somente o comportamento distinto dos hospitais face à complexidade e à gravidade dos casos tratados, bem como a respectiva associação com os respectivos resultados.

Finalmente, como foi referido, pese embora o facto de que os Índices de Casemix não constituírem uma medida directamente caracterizadora das admissões dos hospitais, os mesmos terem sido utilizados para o efeito, nomeadamente quando se associa o perfil das admissões com os resultados dos hospitais.

Esta situação tem razões de dupla natureza:

- Por serem demasiado agregados os Índices de Casemix não permitem uma boa descrição do perfil das admissões hospitalares, pelo que a caracterização da situação utiliza os outros indicadores;

- Pelo seu nível de agregação e porque contêm em si mesmo mecanismos para ponderarem os diversos produtos hospitalares, os Índices de Casemix são mais válidos para análises explicativas.

Estas justificações vão por sua vez de encontro aos argumentos apresentados por Hornbrook (1982) a favor e contra a utilização de índices escalares (a este propósito ver Capítulo Enquadramento Teórico e Questões Práticas).

Para a avaliação do desempenho hospitalar foram utilizados indicadores referentes a três dimensões, a efectividade, a eficiência e o desempenho global, ponderando igualmente as outras duas dimensões.

Para a efectividade escolheu-se um indicador de resultados finais – a mortalidade, tendo sido a validade do mesmo abordada no Capítulo Enquadramento Teórico e Questões Práticas e discutida nos pontos anteriores desta Discussão Metodológica.

Sem pretender esgotar o assunto, deve evidenciar-se que a mortalidade ajustada pelo risco, constitui um bom indicador dos resultados em saúde, exprimindo de uma forma directa as mortes evitáveis ou ainda os ganhos em saúde de cada instituição de saúde (Fink, Yano e Brook, 1989 e lezzoni, 1997b).

No entanto, para além das reservas apontadas por Hofer e Hayward (1996) e por Thomas e Hofer (1998 e 1999) acerca da utilização da mortalidade ajustada pelo risco como indicador da qualidade dos cuidados de saúde, podem ainda referir-se os seguintes aspectos como potenciais problemas:

- Qualidade dos dados utilizados. Este aspecto assume duas perspectivas: o balanço entre os dados administrativos e os dados clínicos (ver Enquadramento Teórico e Questões Práticas e ponto seguinte da Discussão Metodológica – Instrumentos) e a fiabilidade dos dados, tanto decorrente dos próprios problemas da CID-9-CM, nomeadamente a sua precisão e possibilidade de sobreposição (lezzoni, 1990), como dos próprios sistemas de classificação de doentes, como a possibilidade de manipulação (Hornbrook, 1982 e 1985; Thomas, Ashcraft e Zimmerman, 1986; Thomas e Ashcraft, 1989 e lezzoni 1997f);
- A possibilidade de todas as dimensões do risco não estarem devidamente identificadas e medidas, como pode decorrer de situações em que os óbitos tenham ocorrido numa situação precoce do internamento (Zimmerman, 1989; Knaus et al, 1993; lezzoni, 1993 e Zalkind e Estaugh, 1997), as preferências dos consumidores para tratamentos menos agressivos (Murphy e Cluff, 1990 e Philips et al, 2000), a omissão de variáveis, como o estatuto funcional (Murphy e Cluff, 1990; Kazis, 1991; Davis et al, 1995; lezzoni, 1995 e Hornbrook e Goodman, 1996).

Neste sentido, sugere-se a importância da realização de estudos futuros, nomeadamente sobre:

- A investigação entre a associação de medidas de estrutura, processo e resultados para a avaliação da efectividade e da qualidade dos cuidados prestados;
- A comparação entre diversos sistemas de classificação de doentes na avaliação da efectividade dos hospitais, tanto entre sistemas administrativos, com a utilização dos All Patient Refined DRGs ou dos International Refined DRGs, por exemplo, como com sistemas clínicos, dos quais se pode referir os MedisGroups, mas para os quais é necessário alterar o sistema de informação hospitalar (este assunto será abordado no ponto seguinte deste capítulo).

Para a eficiência escolheu-se igualmente um indicador de resultados – a demora média ajustada pelo risco. No Capítulo Enquadramento Teórico e Questões Práticas, foram evidenciados os principais problemas deste indicador, mas sendo igualmente referido que diversos estudos internacionais revelam uma grande associação entre custos médios e demora média.

Contudo, atendendo por um lado, a que estes estudos foram maioritariamente realizados nos EUA e a que, por outro lado, o sistema estrutural, organizativo e de gestão é completamente diferente em Portugal, sendo bons exemplos, a universalidade do sistema, com as naturais consequências em termos de acessibilidade, o tipo de propriedade e de finalidades dos hospitais e ainda a política de gestão de recursos humanos, poderá ser expectável que se observe uma relação diferente entre custos médios e demora média.

Para que tal seja investigado é indispensável a implementação de um sistema de custeio por doente, o qual por sua vez pode posteriormente ser agregado em função da unidade de conta ou produto que se considerar conveniente.

Mais uma vez a literatura internacional evidencia a importância do ajustamento pelo risco e a existência de indicações potencialmente diferentes entre os diversos sistemas de classificação de doentes, mas levanta menos reservas sobre a utilização da demora média como indicador de eficiência (Hughes et al, 1996 e Iezzoni et al, 1996c).

Neste sentido, à semelhança do que foi referido para a efectividade, sugere-se que no futuro sejam realizados estudos comparativos, utilizando igualmente sistemas administrativos e clínicos.

O indicador global de desempenho foi calculado através da média dos resultados dos hospitais para a efectividade e para a eficiência, estando patente que os hospitais são ordenados de forma distinta em função de diferentes pesos específicos para cada uma destas dimensões.

Estes resultados sugerem três questões de natureza diferente: (1) os estudos sobre a avaliação do desempenho dos hospitais devem sempre indicar qual a dimensão e indicadores em estudo, bem como a metodologia para a sua medição; (2) esses mesmos estudos devem sempre indicar qual o peso específico dado a cada dimensão e indicador, visto que os hospitais são classificados de forma diferente em função de ponderações distintas e (3) os estudos sobre a avaliação do desempenho dos hospitais que incluam somente uma dimensão, podem disponibilizar indicações enviesadas ou mesmo erróneas, pelo que não são recomendados.

Assim, sugere-se que no futuro este assunto seja analisado com mais detalhe, essencialmente no que se refere à definição das dimensões a incluir, bem como à respectiva ponderação.

Para o apuramento dos indicadores de efectividade e de eficiência referiu-se que, com base na literatura internacional, a melhor metodologia consistia na comparação entre os valores observados e os valores esperados.

Para o cálculo dos valores esperados da mortalidade e da demora média foram utilizados os “outputs” do “software “ do Disease Staging, sendo cada um destes “outputs” calculado com modelos específicos (MEDSTAT, 2001), o que está de acordo com o discutido em diversos estudos internacionais que referem a necessidade de modelos específicos para se estimar a mortalidade e a demora média (Hornbrook, 1982; Goldfarb, Hornbrook e Higgins, 1983; Coffey e Goldfarb, 1986; Murphy e Cluff, 1990; Knaus et al, 1993; Shwartz et al, 1996; Iezzoni et al, 1996c e Houchens, 2002).

O “software” do Disease Staging para além de definir modelos diferentes para a estimação da mortalidade e da demora média, também utilizou técnicas estatísticas distintas, a regressão logística para a mortalidade e a regressão linear para a duração de internamento (MEDSTAT, 2001), o que mais uma vez está em conformidade com o recomendado para o efeito (Johnson, 1996; Ash e Shwartz, 1997 e Shwartz e Ash, 1997).

Igualmente de acordo com o referido na literatura, os dados originais do Disease Staging não se encontram calibrados para os dados portugueses (Flanders et al, 1999; Justice, Covinsky e Berlin, 1999 e Iezzoni, 1999), pelo que foi necessário proceder à recalibração dos dados.

Tendo em atenção os modelos originais do Disease Staging, a metodologia para a recalibração dos dados foi igualmente distinta, pelo que se passam a discutir separadamente os procedimentos da mortalidade e da demora média.

A recalibração da mortalidade foi feita em separado para os casos cirúrgicos e médicos e foi realizada em duas fases.

No 1º passo, relacionou-se o “outcome” dos doentes (alta vivo = 0 e alta falecido = 1) com as previsões do Disease Staging. Foram realizadas 302 regressões logísticas distintas, das quais 214 para os casos médicos. Não foi possível obter recalibração em 26 situações (doenças), das quais 17 correspondem a casos médicos.

A estas 26 situações corresponderam 472502 episódios de internamento, cerca de 20% da população em estudo. Nos casos cirúrgicos foram encontrados 387498 episódios de internamento (43% dos casos cirúrgicos), enquanto que nos casos médicos não existiu recalibração em 85004 doentes (6% dos casos médicos).

Esta situação é relativamente interessante, visto que embora existam mais episódios sem recalibração nos casos cirúrgicos do que nos casos médicos, a este último grupo de doentes correspondem taxas de mortalidade mais elevadas.

Assim, passou-se ao 2º passo, em que a variável dependente é o “outcome”, com a mesma definição referida anteriormente e a variável explicativa é o valor previsto para a mortalidade do 1º passo.

Foram realizadas 75 regressões logísticas, das quais 50 para os casos médicos, não se alcançando a recalibração em 3 situações, todas dos casos médicos e às quais correspondem 42744 doentes, 1.8% do total de episódios e 3.3% dos casos médicos.

Estes resultados da 2ª fase da recalibração justificam só por si o procedimento, visto que se observou uma redução de 429758 episódios de internamento sem recalibração.

As três situações sem recalibração são Pneumonia por microrganismo não especificado (código 486 da CID-9-MC) – 37906 episódios, Tumor maligno da mama feminina, outras localizações específicas (código 1748) – 2576 episódios e Doenças cerebrovasculares agudas, mal definidas (código 486) – 2262 episódios.

Estes resultados para além de não porem em questão o processo de recalibração, apresentam uma característica interessante – o carácter residual ou agregativo dos diagnósticos sem recalibração.

Estes aspectos chamam a atenção para as seguintes questões:

- Podem indiciar problemas na codificação por parte dos hospitais;
- Podem corresponder a situações de saúde mal estudadas nos hospitais, entre as quais as “Doenças cerebrovasculares agudas, mal definidas” ou a “Pneumonia por microrganismo não especificado” são bons exemplos;

- Podem corresponder a comportamentos diferentes entre os EUA e Portugal;
- Podem indiciar alguns problemas do “software” do Disease Staging para estas situações.

Para se recalibrar a duração de internamento duas estratégias seriam possíveis:

- Realizar regressões lineares por DRG, em que a variável dependente é logaritmo natural da duração de internamento e a variável explicativa é o valor previsto para a duração de internamento do “software” do Disease Staging;
- Calcular a demora média por DRG em função da duração de internamento esperada por doente (“output” do “software” do Disease Staging) ponderada por um valor de referência (demora média observada por DRG).

Conforme foi referido no Capítulo Resultados – Avaliação da Eficiência a duração de internamento é conhecida por ter uma assimetria positiva (Shwartz et al, 1996; lezzoni et al, 1996c; Kulinskaya, Staudte e Hales, 2002 e Clark e Ryan, 2002). No caso português esta realidade é igualmente patente visto que a “skewness” é igual a 14.8 e, após a eliminação dos “outliers” igual a 3.4.

Quando se está na presença de distribuições enviesadas em relação à distribuição normal uma das formas de eliminar este fenómeno é a utilização do logaritmo natural (Johnson, 1996). Mais uma vez em Portugal esta transformação reduziu fortemente este enviesamento, sendo aquela estatística de 0.33 e de 0.21, respectivamente para todos os doentes e para a população excluindo os “outliers”.

Contudo, neste estudo optou-se por uma segunda alternativa para se recalibrar a duração de internamento – a utilização dos valores originais do Disease Staging, ponderados por um valor de referência (ver Capítulo Metodologia).

A preferência por esta metodologia tem dois tipos de justificação, o primeiro de natureza mais operacional, a sua mais fácil aplicação e interpretação e o segundo, porque existe alguma evidência que os resultados das duas estratégias são similares (Hughes, 2002b).

Neste sentido, será interessante discutir se a utilização da primeira estratégia conduziria ou não a resultados distintos. Para se analisar esta situação adoptaram-se os seguintes procedimentos:

- Análises de regressão simples entre:
 - Logaritmo natural da duração de internamento (LnDI) e o logaritmo natural dos pesos relativos dos DRGs (LnPR);

- LnDI e LnDIEsp, em que DIEsp=Duração de Internamento Esperada;
- Análise de regressão múltipla entre LnDi (variável dependente) e LnPR e LnDIEsp (variáveis explicativas).

O coeficiente de determinação (r^2) foi de 0.229 quando se utiliza o peso relativo dos DRGs como variável explicativa, de 0.351 quando se utiliza a duração de internamento prevista pelo Disease Staging e de 0.377, quando se utilizam ambas as variáveis.

Estes resultados permitem desde já evidenciar que as previsões para a duração de internamento do Disease Staging são superiores às disponibilizadas pelos DRGs e que o modelo que utiliza ambas as variáveis é o que apresenta melhor ajustamento, quando este está a ser medido pelo coeficiente de determinação.

Em seguida, estimaram-se os novos valores esperados da duração de internamento segundo o modelo que utiliza o peso relativo dos DRGs e os valores previstos para a duração de internamento do Disease Staging como variáveis explicativas.

Uma primeira conclusão pode ser retirada através da análise de correlação entre os valores previstos quando se utiliza o valor de referência (2ª estratégia) e os resultados do modelo de regressão (1ª estratégia). Como era esperado esta associação é forte e significativa, sendo o coeficiente de correlação de 0.876.

Em seguida serão feitas comparações entre as demoras médias previstas para os DRGs e para os hospitais, pelas duas estratégias, tendo em vista pesquisar se existem ou não diferenças entre as médias.

Para se comparar as demoras médias previstas por DRG pelas duas estratégias referidas, utilizou-se o teste “t” para amostras emparelhadas (o coeficiente de correlação entre as duas previsões foi de 0.822), tendo-se concluído que não existem diferenças significativas entre as duas médias (o intervalo de confiança para a diferença das médias foi de -0.921 a 0.161, sendo a probabilidade de significância da estatística “t” superior a 0.05).

Para a comparação das médias previstas para os hospitais, utilizou-se igualmente o teste “t” para amostras emparelhadas (coeficiente de correlação igual a 0.937), tendo-se concluído que não existem diferenças significativas entre as duas médias (o intervalo de confiança para a diferença das médias foi de -0.001 a 0.208, sendo a probabilidade de significância da estatística “t” superior a 0.05).

Assim, pode concluir-se que todas as indicações utilizadas, coeficientes de correlação para a duração de internamento, para a demora média por DRG e para a demora média por hospital e ainda os testes para amostras

emparelhadas, permitem concluir que as duas estratégias para a recalibração da duração de internamento estão a dar a mesma informação, pelo que, como era esperado, a utilização de um valor de referência para recalibrar a duração de internamento não conduziu a enviesamentos no estudo.

Daqui, pode concluir-se que o processo de recalibração da duração de internamento foi concluído com sucesso, pelo que os resultados para a avaliação da eficiência são consistentes e válidos.

6.1.5. Instrumentos

No Capítulo Enquadramento Teórico e Questões Práticas foram caracterizados os dois sistemas de classificação de doentes utilizados neste estudo – os Diagnosis Related Groups (DRGs) e o Disease Staging.

Foi igualmente referido que enquanto os DRGs, embora tenham sido inicialmente desenhados para a revisão de utilização, têm uma grande aplicação no financiamento dos hospitais, pelo que pretendem construir produtos com consumo de recursos homogêneo, enquanto que o Disease Staging está baseado num modelo biológico, no qual os estadios de gravidade reflectem a progressão da doença em função do respectivo prognóstico (Hornbrook, 1982 e Thomas, Ashcraft e Zimmerman, 1986).

Nesta conformidade, enquanto que os DRGs são um sistema de classificação de doentes profundamente empírico, pelo que estão baseados no que foi feito (Vladeck e Kramer, 1988) o Disease Staging com uma abordagem mais conceptual, está mais preocupado com as características dos doentes e com o que deve ser feito (Markson et al, 1991).

Estes aspectos têm levado a uma enorme polémica, na qual autores como Charbonneau e outros, 1988 e Vladeck e Kramer, 1988 afirmam que, apesar das imperfeições que contêm, os DRGs são o melhor sistema de classificação de doentes existente até ao momento, enquanto que outros como McIlrath, 1985 (citado em Rosko, 1988); Horn, 1986 e Horn e Moses, 1986 e referem que os DRGs não têm significância clínica e que não são facilmente interpretáveis e utilizáveis pelos médicos.

Em relação ao Disease Staging são encontradas afirmações como as produzidas por Charbonneau e outros (1988) “É bastante difícil afirmar que o Disease Staging mede a gravidade”, ou por Aronow (1988) na qual o Disease Staging é considerado como um sistema com pouca capacidade para estabelecer previsões sobre o consumo de recursos, ou ainda por Coffey e Goldfarb (1986), no qual a principal vantagem do Disease Staging reside no facto de não depender do consumo de recursos, mas sim da progressão da doença para definir produtos.

Tendo em atenção estes aspectos as principais críticas apresentadas aos DRGs são as seguintes (Coffey e Godfarb, 1986; Culler e Ehrenfreund, 1986; Aronow 1988; Rosko, 1988; Thomas e Ashcraft, 1989 e Costa, 1994):

- Não consideram a gravidade ou incluem categorias com grandes variações na gravidade;
- Não apresentam significância clínica;

- O consumo de recursos é medido em função da prática e não do que deve ser feito;
- A sua dependência em relação aos resumos de alta e às incorrecções nele contidas;
- Apresentam neutralidade económica, pelo que permitem a escolha de tratamentos que podem ter uma relação menor de custo-efectividade;
- Existe uma grande heterogeneidade entre grupos, essencialmente atribuível a identificação incorrecta dos produtos, a variações na prática clínica, à definição de grupos demasiado agregativos e a grandes variações na gravidade dos doentes;
- Atendendo à sua heterogeneidade clínica não são adequados para se avaliarem os resultados em saúde, designadamente a mortalidade e as readmissões;
- No que se refere ao financiamento também têm sido apresentados alguns problemas, nomeadamente de criar incentivos para se tratarem os casos com preços mais elevados (casos cirúrgicos), à possibilidade de selecção adversa com eliminação dos casos mais graves dentro do mesmo produto e à possibilidade de existência de conflitos com a qualidade dos cuidados prestados.

Em relação ao Disease Staging as principais críticas apontadas dizem respeito (Thomas, Ashcraft e Zimmerman, 1986; Coffey e Goldfarb, 1986; Charbonneau et al, 1988; Rosko, 1988; Costa, 1991):

- À existência de um número de produtos elevado;
- À sua inconsistência para prever consumo de recursos;
- À inadequação do sistema para classificar alguns doentes;
- À inexistência de uma medida global de gravidade;
- À sua dependência em relação aos resumos de alta e às incorrecções nele contidas;
- À inconsistência existente em alguns sub-estádios nos quais a progressão da gravidade não está demonstrada;
- À sobreposição de categorias;
- À impossibilidade de se compararem as diversas doenças;
- No que se refere ao financiamento são igualmente apresentados alguns problemas, designadamente a possibilidade de penalização dos hospitais com procedimentos com custos mais elevados (mais complexos), visto que o preço exprime o custo médio e ao desincentivo para se tratarem casos cirúrgicos.

Como principais vantagens dos DRGs têm sido evidenciados os seguintes aspectos (Hornbrook, 1982; Plomman, 1982; Culler e Ehrenfreund, 1986; Charbonneau et al, 1988; Rosko et al, 1988; Vladeck e Kramer, 1988 e Costa, 1994):

- O sistema resultou num número manejável de grupos;
- O sistema baseia-se em dados disponíveis nos hospitais, o que facilita o nível de comparação e a sua exequibilidade;
- O sistema de classificação é conceptualmente atractivo, porque pretende estabelecer padrões de consumo de recursos, baseando-se nas semelhanças e diferenças entre os doentes;
- Em relação ao financiamento deve evidenciar-se o carácter prospectivo, decorrente do conhecimento prévio dos preços de pagamento, a definição objectiva da unidade de pagamento e o incentivo existente para a prestação de cuidados de saúde mais eficientes.

Para o Disease Staging têm sido referidas as seguintes vantagens (Thomas, Ashcraft e Zimmerman, 1996; Coffey e Goldfarb, 1986; Aronow, 1988 e Costa, 1991):

- O carácter atractivo do sistema, visto que se baseia na progressão biológica da doença, na qual os estadios, para além de serem homogéneos entre si, representam ainda um maior risco de morbilidade e de mortalidade à medida que vão aumentando;
- A adequação do sistema para se estabelecerem previsões sobre os resultados de saúde, nomeadamente a mortalidade;
- O sistema demonstra ainda um consumo de recursos diferente por estadio de gravidade;
- O sistema, visto que se baseia no que deve ser feito, permite a identificação dos recursos necessários, bem como a avaliação dos recursos utilizados;

Por outro lado, num estudo realizado por Thomas, Ashcraft e Zimmerman (1986), onde se comparava o desempenho de diversos sistemas de medição da severidade do estado do doente, os autores embora referindo que os PMCs e os MedisGroups apresentam maior validade de construção, que o APACHE apresenta maior validade de conteúdo, que os PMCs e o Disease Staging apresentam maior validade de previsão, que o APACHE e os MedisGroups apresentam maior fiabilidade, que o APACHE tem menor probabilidade de manipulação e que os PMCs e o Disease Staging têm menores custos de implementação e de exploração, acabam por recomendar aos hospitais a utilização dos Patient Management Categories e do Disease Staging.

Noutro estudo realizado por Thomas e Ashcraft (1989), onde se pretendia avaliar a fiabilidade de diversos sistemas de classificação de doentes, os autores concluem que o APACHE, os MedisGroups e os DRGs apresentam uma boa fiabilidade, enquanto que o Disease Staging apresenta uma fiabilidade razoável a boa.

Os autores referem que o desempenho do Disease Staging se deve à utilização dos dados dos resumos de alta, pelo que estranham que os resultados dos DRGs tenham sido diferentes, visto que utilizam a mesma fonte de dados (Thomas e Ashcraft, 1989).

No entanto, acabam por concluir que, aparentemente, a característica dos DRGs que tem conduzido a várias críticas acerca da heterogeneidade clínica, em que vários diagnósticos, por vezes não relacionados entre si, são agregados na mesma categoria, proporciona igualmente uma maior fiabilidade em relação a sistemas que utilizam a mesma base de dados (Thomas e Ashcraft, 1989).

Por outro lado existe ainda uma enorme variedade de estudos que pretendem avaliar o desempenho dos sistemas de classificação de doentes em relação à sua capacidade para estabelecerem explicações e previsões sobre o consumo de recursos e sobre a mortalidade.

Em 1980, Bentley e Butler, defendem que, atendendo às suas características, os Patient Management Categories (PMCs) e o Disease Staging podem ser considerados como boas alternativas aos DRGs.

Em 1982, Ament e outros, ao avaliar o desempenho de três sistemas de classificação de doentes, acabam por concluir que os DRGs e o Disease Staging apresentam o mesmo poder explicativo sobre os custos.

Em 1984, Conklin e outros, quando realizaram um estudo comparativo entre os DRGs e o Disease Staging, encontraram uma relação directa entre consumo de recursos e estadios da gravidade.

Em 1987, Calore e lezzoni concluem que embora os PMCs e o Disease Staging não sejam alternativas aos DRGs, podem ser úteis para modificar alguns dos aspectos dos DRGs.

Neste mesmo estudo, Calore e lezzoni (1987), embora afirmem que o Disease Staging sozinho não aumenta o poder explicativo dos DRGs e que a utilização conjunta destes dois sistemas de classificação apresenta um melhor ajustamento, acabam por recomendar a utilização isolada dos DRGs.

Em, 1988, Charbonneau e outros, afirmam que a utilização isolada do Disease Staging apresenta reduções nas variações da duração de internamento menores que as produzidas pelos DRGs. Atendendo a que a utilização conjunta destes dois indicadores introduz poucas alterações no cenário avaliativo dos DRGs, acabam por sugerir que a utilização do Disease Staging é inútil.

Diversos estudos, como os realizados por lezzoni e outros (1995a), lezzoni e outros (1996a), lezzoni e outros (1996b), lezzoni e outros (1996d), lezzoni e outros (1996e) e lezzoni (1997c), concluem que os All Patient Refined Diagnosis

Related Groups (APRDRGs) e o Disease Staging têm uma calibração e discriminação praticamente iguais (embora com uma ligeira supremacia do Disease Staging) e que existe uma grande concordância entre estes dois sistemas de classificação de doentes para avaliarem o desempenho dos hospitais. Quase todos estes estudos demonstram no entanto que os resultados na avaliação do desempenho dos hospitais são diferentes quando se utilizam os modelos clínicos, como os MedisGroups ou o APACHE, por exemplo.

Em relação à duração de internamento podem referir-se dois estudos realizados em 1996, por Schwartz e outros e por Iezzoni e outros (1996c), ambos avaliando a importância da severidade nas variações da demora média, estando o primeiro aplicado à “Fratura de anca” e o segundo à “Pneumonia”, os autores concluem que os APRDRGs e o Disease Staging concordam na avaliação do desempenho dos hospitais, embora os APRDRGs apresentem um poder explicativo ligeiramente superior ao do Disease Staging no caso da “Fratura da anca” e praticamente igual para a “Pneumonia”. Nestes estudos, admitindo-se a existência de uma relação não linear entre a duração de internamento e a gravidade foi introduzido o termo ao quadrado desta variável.

Tendo em atenção os aspectos anteriormente referidos e atendendo igualmente a que, como foi referido, neste estudo estão a ser utilizados os DRGs e o Disease Staging, em seguida serão discutidos os seguintes aspectos:

- Vantagens e inconvenientes dos sistemas de classificação utilizados;
- Validade dos sistemas utilizados;
- Recomendações e passos futuros.

Em relação aos DRGs as vantagens e inconvenientes anteriormente apresentados mantêm-se com a versão que está a ser utilizada. No entanto, as versões que contêm a gravidade (ver Capítulo Metodologia – Instrumentos) podem obviar a alguns dos problemas referidos. Voltar-se-á a esta temática no final deste ponto.

Em relação ao Disease Staging algumas das críticas referidas estão completamente ultrapassadas com a versão que está a ser utilizada neste estudo, nomeadamente no que se refere:

- À sua inconsistência para prever consumo de recursos. De facto, o “software” do Disease Staging estabelece um modelo específico para a previsão do consumo de recursos, tanto em termos de custos, como da duração de internamento. Para a duração de internamento o coeficiente de determinação (r^2) foi de 0.36 (MEDSTAT, 2001). Para os dados portugueses esta estatística foi de 0.35, enquanto que a utilização do peso específico dos DRGs foi de 0.23. Assim, pode concluir-se que o Disease Staging demonstra alguma capacidade para prever o consumo

de recursos (a este propósito ver igualmente artigos de Schwartz e outros (1996) e Iezzoni e outros (1996c);

- À inexistência de uma medida global de gravidade. De facto, as versões original e codificada atribuem estadios e sub-estadios para o diagnóstico principal (incluindo as complicações) e para as comorbilidades, pelo que esta crítica parece ter cabimento. Contudo, mais uma vez o “software” do Disease Staging parece estar a contribuir para se ultrapassar este problema, visto que a mortalidade prevista, desde que apresente discriminação, pode ser utilizada como representativa da gravidade de cada doente;
- À impossibilidade de se compararem doenças. Mais uma vez os valores previstos para o consumo de recursos e para a mortalidade, desde que devidamente ajustados, contribuem para obviar a esta questão, ultrapassando os problemas originais do Disease Staging, no qual os estadios de gravidade não representam o mesmo para diferentes doenças;
- À existência de um número elevado de categorias. Esta questão pode constituir um problema sobretudo estatístico, visto que podem existir células (produtos) com poucas observações. Quando o problema é deste tipo, a utilização das escalas da mortalidade e do consumo de recursos (derivadas dos valores previstos) pode concorrer para contornar a situação, visto que constituem índices escalares, mantendo-se no entanto as críticas referidas para este tipo de abordagem;
- À inadequação do sistema para classificar alguns doentes e à sobreposição de categorias. Mais uma vez, o desenvolvimento do Disease Staging está a contribuir para reduzir ou mesmo eliminar este problema, visto que no momento das críticas existiam cerca de 400 doenças identificadas e com estadios, enquanto que na versão actual estão consideradas cerca de 650 doenças, pelo que se considera que estes problemas estão minimizados.

No que respeita à validade dos sistemas, para a validade de construção e de conteúdo referem-se diversos estudos internacionais para os DRGs (Hornbrook, 1982 e 1985; Charbonneau et al, 1988; Rosko, 1988 e Vladeck e Kramer, 1988) e para o Disease Staging (Hornbrook, 1982 e 1985; Thomas, Ashcraft e Zimmerman, 1986 e Aronow, 1988).

No que se refere à validade de previsão para além dos resultados de diversos estudos anteriormente apresentados, deve evidenciar-se o seguinte:

- Para a mortalidade, pelos motivos referidos, somente se pode utilizar o Disease Staging. O processo de recalibração dos valores esperados aos dados portugueses, o qual foi concluído com sucesso, com excepção dos três diagnósticos anteriormente referidos, permite igualmente concluir pela validade de previsão deste instrumento;

- Para o consumo de recursos, foi igualmente referido o poder explicativo dos DRGs, do Disease Staging e da utilização conjunta, sendo o coeficiente de determinação (r^2) de 0.229, de 0.351 e de 0.777, respectivamente. Estes resultados, em função dos encontrados em diversos estudos internacionais devem ser considerados positivos, pese embora o facto de se esperar um “ r^2 ” superior a 0.50 para existir um bom ajustamento;
- A este propósito somente mais um aspecto relevante – a consideração de modelos específicos para a previsão da duração de internamento. De facto, quando se utiliza directamente o valor esperado da mortalidade após recalibração aos dados portugueses o “ r^2 ” foi somente de 0.073 e de 0.122 quando se utiliza o termo ao quadrado (o que significa que as durações de internamento são mais reduzidas em níveis de gravidade menos e mais elevados, nesta última situação por causa dos mortos). A introdução do peso relativo dos DRGs aumenta o “ r^2 ” para 0.266. Foi muito provavelmente por esta situação, utilização dos valores esperados da mortalidade do Disease Staging em detrimento dos valores estimados directamente para a duração de internamento, que as conclusões dos estudos de Shwartz e outros (1986) e de Iezzoni e outros (1996c) referem a pouca influência da gravidade para explicar a duração de internamento.

Assim, para a realização de trabalhos futuros recomenda-se a utilização de sistemas de classificação de doentes que utilizem dados clínicos e ainda a consideração das versões dos DRGs que incluem a severidade, designadamente os International Refined ou os All Patient Refined.

Este aspecto, para além de permitir a comparação entre os diversos sistemas de classificação de doentes, permitirá igualmente a existência de um cenário mais válido para avaliar o desempenho dos hospitais.

6.1.6. Fontes de Informação

No Capítulo Revisão da Literatura – Modelos de Ajustamento pelo Risco foi caracterizada a situação, no que respeita ao balanço entre os sistemas que utilizam dados clínicos e os que usam os dados administrativos.

Como foi igualmente referido, neste estudo são utilizados dois sistemas de classificação de doentes que utilizam dados administrativos (os DRGs e o Disease Staging) simplesmente pela razão de no sistema de informação de rotina dos hospitais portugueses não ser possível utilizar elementos clínicos.

No ponto anterior foi igualmente referido que por vezes, essencialmente na avaliação da mortalidade, os sistemas clínicos disponibilizam informações diferentes dos sistemas administrativos, pelo que algumas reservas devem ser tomadas em consideração quando se avalia o desempenho dos hospitais utilizando dados administrativos, pese embora o facto de Polanczyk e outros (1998) referirem que a utilização de dados administrativos para se estabelecerem previsões sobre a mortalidade constitua uma metodologia adequada.

Os estudos de Aronow (1988), de lezzoni (1997c) e de Wray e outros (1997) sintetizam bem o problema:

- Os sistemas de classificação de doentes que utilizam os dados administrativos são mais fáceis de implementar e de gerir, pelo que os custos associados são menores;
- Estes sistemas, por sua vez, podem revelar algumas insuficiências na significância clínica, devido à ambiguidade, sobreposição e falta de sensibilidade para os procedimentos existentes na CID-9-MC, classificação que está na origem daqueles sistemas (Hornbrook, 1982; Mullin, 1985 e lezzoni e Moskowitz, 1986);
- Outros problemas da CID-9-MC dizem respeito à limitação da sua informação clínica, visto que somente alguns códigos contêm informação sobre a gravidade (o código 799.1 – “Paragem respiratória” é um bom exemplo) ou sobre as causas para que a situação tenha acontecido (o código 995.4 – “Choque devido a anestesia” é igualmente um bom exemplo), pelo que não se podem atribuir ao doente ou ao prestador as causas do problema de saúde;
- Mesmo entre os sistemas que utilizam os dados administrativos os problemas de significância clínica estão mais patentes nos DRGs, do que naqueles que se baseiam na gravidade, como o Disease Staging;
- Os sistemas administrativos revelam menos precisão na identificação das comorbilidades e das complicações (Hartz et al, 1994; lezzoni et al, 1994b e 1994c; McCarthy et al, 2000; Powel, Lim e Heller, 2001; Quan, Parson e Ghali, 2002; Romano et al, 2002 e Geraci, 2002),

embora o Disease Staging revele um desempenho superior quando comparado com os DRGs;

- Finalmente, os sistemas administrativos utilizam dados de toda a estadia, não identificando a situação do doente no momento de admissão e outras adquiridas durante a estadia, pelo que podem conter problemas para a avaliação da qualidade dos cuidados prestados. Mesmo para o consumo de recursos, atendendo a que os médicos tratam os doentes em função dos sinais vitais e dos sintomas presentes no momento de admissão, os quais por sua vez se alteram durante a estadia e que consequentemente condicionam o consumo de recursos, o que dificilmente é captado por sistemas que atribuem retrospectivamente diagnósticos, pelo que podem existir variações não justificadas.

É por estas razões que se defende que a utilização de sistemas clínicos que disponibilizem informações sobre a situação do doente no momento da admissão, durante a estadia e na alta são mais adequados para a avaliação do desempenho.

Mais uma vez, se recorda que este procedimento não foi seguido neste estudo devido a limitações no sistema de informação. Contudo, tal não significa que o ajustamento pelo risco, perde relevância ou validade, visto que produz análises mais sólidas que a avaliação pelos resultados brutos.

No entanto, reforça-se a necessidade de se alterar o sistema de informação para conter dados clínicos, para que uma análise mais detalhada do Disease Staging (critérios clínicos) ou dos MedisGroups, por exemplo, pode constituir um elemento facilitador.

Finalmente, deve referir-se o problema da codificação de diagnósticos principais e essencialmente secundários. Segundo lezzoni e outros (1992b) a questão assume duas perspectivas: a qualidade no processo de codificação e a que uma maior precisão pode ser alcançada com a utilização de um maior número de diagnósticos secundários.

Em relação ao segundo aspecto – maior precisão do sistema de ajustamento pelo risco – parece genericamente comprovada com a análise temporal dos valores da taxa de mortalidade observada e da taxa de mortalidade esperada (ver Metodologia – Recalibração da Mortalidade Esperada).

De facto, para além dos valores da mortalidade esperada aumentarem sucessivamente entre 1999 e 2001, neste último ano, o único em que existe a possibilidade de codificar até 19 diagnósticos secundários (complicações e comorbilidades), encontra-se a menor razão entre a taxa de mortalidade observada e a taxa de mortalidade esperada.

Atendendo à importância do assunto, sugere-se que no futuro se realizem estudos para melhor qualificar o fenómeno.

Em relação à qualidade da codificação, são conhecidos os esforços recentes para se melhorar este aspecto, através da realização de estudos de fiabilidade (também designados por estudos de auditoria). No entanto, é completamente impossível, pelo menos neste momento, proceder a uma associação entre aqueles estudos e os resultados aqui apresentados.

A única análise que se pode fazer é meramente quantitativa, ou seja, pesquisar a associação entre o número médio de diagnósticos secundários por hospital e os seus ganhos em termos de efectividade e de eficiência.

Para a efectividade o coeficiente de correlação entre o número médio de diagnósticos secundários e os resultados do hospital depois de ajustados pelo risco é de -0.167 pelo método directo ("z score") e de -0.019, quando se utiliza o método indirecto (taxa de mortalidade padronizada).

Para a eficiência o coeficiente de correlação entre o número médio dos diagnósticos secundários e os resultados ajustados pelo risco é de -0.107 no método directo e de -0.180 no método indirecto.

Estes valores demonstram que não existe nenhuma associação entre o maior esforço de codificação e o desempenho dos hospitais, tanto para a efectividade, como para a eficiência.

Ou seja, quando se considera que o maior volume de diagnósticos secundários possa ser considerado como um "proxy", embora grosseiro, da qualidade de codificação, não parece existir nenhuma relação entre a prática da codificação e o desempenho dos hospitais.

Contudo, para evitar o fenómeno conhecido por "death code creep" (Iezzoni, 1997c), isto é que os hospitais tendam a "sobrecodificar" para melhorarem os seus resultados de saúde, devem ser tomadas medidas concretas para melhorar e normalizar as práticas de codificação dos hospitais, as quais necessariamente devem igualmente passar pela publicitação dos resultados destes estudos de fiabilidade.

Como é lógico, estes esforços embora não resolvam todos os problemas referidos para os sistemas de classificação de doentes que trabalham com os dados administrativos, podem, pelo menos, contribuir para minimizar alguns dos problemas e críticas apontadas a esta abordagem (Tully e Rullon, 2000).

No entanto, reitera-se que para além destes esforços sobre a prática de codificação dos hospitais, é necessário alterar o sistema de informação hospitalar, tendo em vista a criação de bases de dados clínicas e uma maior normalização dos procedimentos hospitalares (Iezzoni, 1997f).

6.2. Discussão dos Resultados

6.2.1. Produção

À semelhança do capítulo Resultados a discussão sobre a produção incidirá sobre as dimensões medição da produção e perfil das admissões.

6.2.1.1. Medição da Produção

No capítulo sobre a caracterização da produção hospitalar observou-se que os dois sistemas de classificação de doentes em estudo – os Diagnosis Related Groups e o Disease Staging estão a dar informação completamente diferente sobre a produção dos hospitais, quando esta é medida pelo Índice de Casemix.

Esta conclusão é válida para qualquer das perspectivas de análise consideradas, com as seguintes particularidades:

- O Disease Staging identifica maiores diferenças no “casemix” dos hospitais, quando estes são identificados pelo respectivo Tipo;
- Alteração no “casemix” regional em função do respectivo sistema de classificação de doentes, embora com uma atenuação na razão entre valores extremos, quando comparada com a existente por Tipo de hospital;
- Identificação e ordenação completamente distinta por hospital em função do respectivo sistema de classificação de doentes, inclusivamente com hospitais que passam do “top ten” do Índice de Casemix calculado pelos DRGs, para os hospitais com o Índice de Casemix mais baixo, quando calculado pelo Disease Staging;
- Valorização distinta entre os DRGs e o Disease Staging para os casos cirúrgicos e médicos, com o primeiro sistema de classificação de doentes a apurar Índices de Casemix mais elevados para os casos cirúrgicos e o Disease Staging para os casos médicos;
- Valorização completamente distinta do Índice de Casemix por Agrupamento de Doenças;
- Menores diferenças, tanto por gravidade dos doentes, como por idade nos DRGs do que no Disease Staging. Ou seja, os Diagnosis Related Groups admitem uma maior homogeneidade entre doentes para estas características do que o Disease Staging.

Tendo em atenção estes aspectos, em seguida irão ser analisadas e discutidas as principais razões explicativas das diferenças encontradas.

Neste sentido, a questão que se coloca de seguida é investigar se estas diferenças entre os dois sistemas de classificação de doentes na atribuição de Índices de Casemix é aleatória, ou seja se não se consegue descobrir um

padrão. Ou, se antes pelo contrário existe um padrão e/ou razões explicativas para estas diferenças.

Para tal foi criado um modelo explicativo por hospital com as seguintes variáveis:

- Produção – é medida pelo número de doentes tratados em cada hospital;
- Sexo – utiliza-se a percentagem de doentes do sexo masculino por hospital;
- Idade – utiliza-se a idade média por hospital;
- Tipo de tratamento – utiliza-se a percentagem de doentes com tratamento cirúrgico por hospital;
- Gravidade – utiliza-se a percentagem de doentes tratados no escalão de severidade 1 (doentes menos graves) por hospital;
- Tipo de hospital – utiliza-se uma “dummy” representativa do Tipo de hospital;
- Região de Saúde – utiliza-se uma “dummy” representativa das cinco Regiões de Saúde.

Vão ser efectuadas três análises de regressão múltipla. A primeira em que a variável dependente é o Índice de Casemix por hospital calculado pelos Diagnosis Related Groups. Na segunda a variável dependente é o Índice de Casemix calculado pelo Disease Staging. Na terceira análise a variável dependente é o Índice de Casemix Composto.

Para o Índice de Casemix dos DRGs (ICMDRGs) os principais resultados do Modelo são os seguintes:

R2 ajustado – 0.803

Estatística “F” – 26.248 (significativa a 0.001)

Figura 56
Resultados do Modelo: ICMDRGs

Variável	Coeficiente	Estatística “t”	Significância
Constante	1.271	4.014 *	0.000
Tipo Tratamento	0.695	4.124 *	0.000
Gravidade	-1.408	-4.353 *	0.000
Sexo	0.493	2.263 *	0.027
Produção	-1.097E-06	-1.245	0.217
Idade	0.006	2.539 *	0.013
Centrais e Universitários	0.312	6.068 *	0.000
Especializados	0.172	2.023 *	0.047
Distritais de Nível 1	-0.130	-2.859 *	0.006
IPOs	0.889	10.124 *	0.000
Alentejo	-0.079	-1.137	0.259
Algarve	-0.002	-0.019	0.985
Centro	-0.081	-1.966	0.053
Norte	-0.010	-0.244	0.807

* - valores significativos a 5%

Como principais comentários podem fazer-se os seguintes:

- O nível de produção não tem efeitos significativos sobre o Índice de Casemix. Mesmo assim a relação é negativa, ou seja, quanto maior a produção menor o Índice de Casemix;
- A localização dos hospitais, medida pela Região de Saúde onde estão incluídos, não tem efeitos sobre o Índice de Casemix;
- Para o Tipo de hospital a situação é diferente, sendo todos os coeficientes significativos, com a particularidade de a associação ser directa para os hospitais Centrais, Especializados e IPOs e inversa para os hospitais Distritais de Nível 1;
- Para o Tipo de Tratamento, Gravidade dos Doentes, Idade e Sexo, todos os coeficientes de regressão são significativos. Com a particularidade de quanto mais doentes com tratamento cirúrgico, do sexo masculino e com mais idade, maior é o Índice de Casemix e que quanto maior é a percentagem de doentes menos graves menor é o Índice de Casemix.

Para o Índice de Casemix do Disease Staging (ICMDS) os principais resultados do Modelo são os seguintes:

R2 ajustado – 0.843

Estatística “F” – 36.874 (significativa a 0.001)

Figura 57
Resultados do Modelo: ICMDS

Variável	Coeficiente	Estatística “t”	Significância
Constante	1.933	3.990 *	0.000
Tipo Tratamento	-1.603	-6.217 *	0.000
Gravidade	-1.892	-3.825 *	0.000
Sexo	0.530	1.590	0.116
Produção	-1.39E-06	-1.044	0.300
Idade	1.717E-02	4.805 *	0.000
Centrais e Universitários	0.159	2.021 *	0.047
Especializados	0.124	0.957	0.342
Distritais de Nível 1	0.173	2.489 *	0.015
IPOs	0.428	3.185 *	0.002
Alentejo	-2.00E-02	-0.188	0.852
Algarve	2.98E-02	0.226	0.822
Centro	-0.122	-1.922	0.058
Norte	-0.142	-2.208 *	0.030

* - valores significativos a 5%

Como principais comentários podem fazer-se os seguintes:

- O nível de produção não tem efeitos significativos sobre o Índice de Casemix. Mesmo assim a relação é negativa, ou seja, quanto maior a produção menor o Índice de Casemix;
- A localização dos hospitais, medida pela Região de Saúde onde estão incluídos, parece ter efeitos reduzidos no Índice de Casemix, sendo estatisticamente significativa somente para a Região Norte;
- Para o Tipo de hospital a situação é diferente, sendo significativos os coeficientes de regressão para os hospitais Centrais, os hospitais Distritais de Nível 1 e para os IPOs. Todas associações são directas;
- O sexo não tem igualmente efeitos significativos sobre o Índice de Casemix, embora com uma relação directa;
- Para o Tipo de Tratamento, Gravidade dos Doentes e Idade, todos os coeficientes de regressão são significativos. Com a particularidade de quanto maior a idade dos doentes, maior é o Índice de Casemix e que quanto maior é a percentagem de doentes menos graves e de doentes com tratamento cirúrgico menor é o Índice de Casemix.

A comparação destes dois modelos parece evidenciar a existência de algumas razões justificativas para as diferenças encontradas entre os dois sistemas de classificação de doentes para o apuramento dos Índices de Casemix:

- Em primeiro lugar, o sexo dos doentes não é relevante no Disease Staging, ao contrário do que ocorre nos DRGs;
- A relevância do Tipo de hospital é igualmente distinta, com particular importância nos hospitais Distritais de Nível 1;
- O Tipo de Tratamento é muito marcante nas diferenças entre os dois sistemas de classificação. Enquanto que os DRGs atribuem Índices de Casemix mais elevados a hospitais com mais doentes com tratamento cirúrgico, no Disease Staging a relação é precisamente a contrária, com Índices de Casemix mais elevados a ocorrerem quando existem menos doentes com tratamento cirúrgico.

De facto, atentos os dados referidos, este último aspecto – Tipo de Tratamento – parece ser o factor mais decisivo para o desempenho distinto dos dois sistemas de classificação de doentes.

Este aspecto pode estar associado à própria lógica de construção dos dois sistemas de classificação de doentes em análise. De facto, enquanto que os DRGs se baseiam na identificação de produtos com semelhanças no consumo de recursos e cuja principal finalidade é o financiamento dos hospitais, o Disease Staging centra a sua lógica de desenvolvimento no modelo biológico das doenças, com as naturais vantagens em termos de avaliação da qualidade dos cuidados prestados.

À semelhança do que foi efectuado para os Índices de Casemix calculados pelos DRGs e pelo Disease Staging, procedeu-se a uma análise de regressão para o Índice de Casemix Composto (ICMCom) com as mesmas variáveis explicativas.

Os resultados são os seguintes:

R2 ajustado – 0.864

Estatística “F” – 43.425 (significativa a 0.001)

Figura 58
Resultados do Modelo: ICMCom

Variável	Coeficiente	Estatística “t”	Significância
Constante	1.534	5.341 *	0.000
Tipo Tratamento	-0.491	-3.213 *	0.002
Gravidade	-1.575	-5.369 *	0.000
Sexo	0.485	2.455 *	0.016
Produção	-1.18E-06	-1.494	0.140
Idade	0.011	5.302 *	0.000
Centrais e Universitários	0.219	4.691 *	0.000
Especializados	0.139	1.803	0.075
Distritais de Nível 1	0.029	0.692	0.491
IPOs	0.610	7.670 *	0.000
Alentejo	-0.045	-0.718	0.475
Algarve	0.014	0.181	0.857
Centro	-0.097	-2.589 *	0.012
Norte	-0.075	-1.983	0.051

* - valores significativos a 5%

Como principais comentários podem fazer-se os seguintes:

- O nível de produção não tem efeitos significativos sobre o Índice de Casemix, embora a relação seja negativa;
- A localização dos hospitais, medida pela Região de Saúde onde estão localizados, tem efeitos reduzidos sobre o Índice de Casemix, visto que o coeficiente de regressão somente é significativo para a Região Centro;
- O Tipo de hospital é relevante para o apuramento do Índice de Casemix, sendo sempre significativo e directo, com excepção dos hospitais de Nível 1;
- Para o Tipo de Tratamento, Gravidade dos Doentes, Idade e Sexo, todos os coeficientes de regressão são significativos. Com a particularidade de quanto maior a idade e de doentes do sexo masculino maior é o Índice de Casemix e de que quanto maior é a percentagem de doentes menos graves e a percentagem de doentes cirúrgicos menor é o Índice de Casemix.

Como seria de esperar estes resultados diferem, tanto dos obtidos para o Índice de Casemix dos DRGs, como para o do Disease Staging.

Para o primeiro, as principais diferenças são encontradas no Tipo de Tratamento com inversão do poder explicativo dos casos cirúrgicos, enquanto que quando se compara com o Disease Staging estas diferenças são encontradas na influência do sexo e no comportamento por Tipo de hospital.

6.2.1.2. Perfil das Admissões

No capítulo da apresentação dos resultados para qualificar o perfil das admissões hospitalares foram utilizados três tipos de indicadores – a concentração/diversificação da produção, a complexidade dos casos e a gravidade dos casos.

No que se refere à concentração da produção foram utilizados dois indicadores, um relativo à raridade dos casos em cada hospital (percentagem de produtos diferentes em cada hospital) e o segundo respeitante ao nível de concentração da produção por hospital (percentagem de produtos responsáveis por 80% da quantidade produzida).

Para qualquer destas dimensões recorreu-se aos dois sistemas de classificação de doentes utilizados neste estudo – os Diagnosis Related Groups e o Disease Staging.

No que se refere ao efeito raridade observou-se que, pese embora o facto de o Disease Staging apresentar um número de produtos por hospital diferente do disponibilizado pelos DRGs, os dois sistemas de classificação de doentes estão genericamente a facultar a mesma informação para qualificar o efeito raridade na produção do hospital.

Observou-se igualmente que existe um comportamento heterogéneo entre os hospitais, perspectivando-se a dimensão dos hospitais, medida pela quantidade produzida, como um dos factores que mais contribuem para as diferenças detectadas. Inclusivamente, foi referido que o efeito raridade da produção varia proporcionalmente com o aumento da produção.

Quando se analisa a concentração da produção medida pelo número de produtos diferentes responsáveis por 80% da produção em cada hospital observou-se o seguinte:

- Os DRGs e o Disease Staging estão genericamente a dar a mesma informação sobre a concentração da produção nos hospitais, embora este último sistema de classificação de doentes apresente níveis de concentração mais elevados;

- A existência de um comportamento heterogéneo entre hospitais, mas mais atenuado do que o presente no *efeito raridade*;
- Importância do volume de produção sobre a respectiva concentração, tanto presente com os DRGs, como com o Disease Staging, embora de uma forma mais intensa no primeiro sistema de classificação.

Para a complexidade e a gravidade dos casos tratados, utilizou-se a informação disponibilizada pelos DRGs para se caracterizar o primeiro atributo e pelo Disease Staging para se apurar a gravidade dos doentes.

Em termos gerais constatou-se que a complexidade e a gravidade dos casos disponibilizam informações diferentes sobre a actividade dos hospitais, sendo os casos cirúrgicos mais complexos e os casos médicos mais graves.

Quando se analisam somente os casos cirúrgicos observou-se que, embora os dois sistemas de classificação de doentes concordem que existem mais doentes tratados nos níveis de complexidade e gravidade mais baixos e ainda menos doentes tratados nos graus mais elevados, o Disease Staging identifica maiores diferenças entre a gravidade do que as apuradas para a complexidade dos casos.

Para os casos médicos a situação é quase igual à descrita anteriormente, somente com a particularidade de existirem maiores diferenças entre complexidade do que as observadas na gravidade.

A análise mais desagregada por hospital identificou as seguintes situações:

- Existem diferenças no comportamento dos hospitais no que se refere à complexidade dos casos tratados, as quais por sua vez são mais intensas nos casos mais complexos (aspecto presente para todas as admissões e para os casos cirúrgicos e médicos);
- Os mesmos hospitais apresentam um comportamento diferente entre os casos médicos e os casos cirúrgicos;
- Para a gravidade dos casos a situação é praticamente idêntica à referida para a complexidade, ou seja maiores diferenças entre hospitais nos casos mais graves, embora com uma maior atenuação nos casos médicos;
- A principal diferença entre estas duas perspectivas é encontrada nos hospitais com maiores e menores complexidade e gravidade, visto que são os hospitais do Tipo I os que tratam doentes mais complexos, enquanto que para os casos mais graves tal ocorre nos hospitais do Tipo IV.

Este último aspecto parece indiciar a existência de informações e de comportamentos distintos por parte dos hospitais para a complexidade e para a gravidade dos casos tratados.

Na realidade este fenómeno foi confirmado, essencialmente nos seguintes aspectos:

- Para todas as admissões, os hospitais com maiores diferenças entre a complexidade e a gravidade dos casos são do Tipo I e V, enquanto que com menores diferenças são identificados os hospitais do Tipo IV e III. Observou-se ainda que não existe associação entre a complexidade e a gravidade dos casos tratados e que a gravidade varia inversamente com o aumento da produção, enquanto que a complexidade e a diferença entre a complexidade e a gravidade dos casos varia directamente com o aumento da produção;
- Para os casos cirúrgicos a situação é diferente da descrita anteriormente, visto que existe uma associação inversa entre a complexidade e a gravidade dos casos tratados, embora a associação entre o volume da produção e a complexidade, gravidade e respectiva diferença seja a mesma;
- Nos casos médicos, mais uma vez são identificados comportamentos distintos, designadamente no que se refere a uma associação directa entre a complexidade e a gravidade dos casos tratados e na inexistência de uma relação entre volume de produção e a complexidade dos casos tratados. Contudo, deve referir-se que se mantém a relação inversa entre a produção e a gravidade dos casos tratados e a relação directa entre a produção e a diferença entre complexidade e gravidade.

Atendendo a que este aspecto não é estrutural no panorama hospitalar português, visto que uma análise por doença permitiu concluir por uma relação directa entre a complexidade e a gravidade dos casos (ver capítulo Resultados), apuraram-se de seguida algumas das vantagens dos hospitais derivadas do seu perfil de admissões.

Os principais resultados foram os seguintes:

- A taxa de mortalidade e a demora média variam directamente com a complexidade e com a gravidade dos doentes tratados por hospital, tanto considerados isoladamente, como conjuntamente. Observou-se igualmente que quanto maior a diferença entre a complexidade dos casos e gravidade dos casos tratados menores são a taxa de mortalidade e a demora média;
- O Índice de Casemix dos hospitais medido pelos DRGs é indiferente à gravidade dos doentes, enquanto que os Índices de Casemix do Disease Staging e Composto variam directamente com a complexidade e a gravidade.

Estes resultados evidenciam uma enorme variedade de situações anómalas ou, no mínimo, não esperadas.

Até ao presente momento, a informação existente em Portugal sobre a actividade dos hospitais utilizava a informação constante dos resumos de alta dos hospitais e com recurso a um sistema de classificação de doentes – os Diagnosis Related Groups.

Em função desta informação os hospitais apresentavam um comportamento heterogéneo em relação ao seu perfil de admissões.

No entanto, este resultado aparentemente não conflituava com a universalidade do Serviço Nacional de Saúde, visto que os hospitais com maior Índice de Casemix e portanto os que tratavam casos mais complexos eram simultaneamente aqueles que produziam mais.

Atentas igualmente as especificidades da organização da oferta em Portugal é natural deduzir que a hospitais com maior produção correspondem organizações de saúde com maior diferenciação, normalmente qualificados como hospitais universitários, centrais e especializados, como é o caso dos Institutos de Oncologia.

Ainda dentro desta linha de pensamento deve igualmente referir-se que os hospitais que tratavam casos menos complexos apresentavam resultados brutos, nomeadamente a taxa de mortalidade e a demora média, melhores que os hospitais que tratavam doentes mais complexos.

Estas premissas e resultados eram frequentemente utilizados como justificativos dos diferentes desempenhos dos hospitais, argumentando-se ainda frequentemente que os hospitais de maior dimensão produtiva, por estarem no fim da linha hierárquica de referência, seriam os que teoricamente teriam menos possibilidade de escolher doentes e concomitantemente tratariam os doentes mais graves.

As consequências naturais deste fenómeno eram de dupla natureza. Em primeiro lugar, argumentava-se que seria muito difícil comparar a actividade dos hospitais, visto que os mesmos tratam doentes completamente diferentes e, em segundo lugar, a maior complexidade e gravidade dos casos tratados era, só por si, justificativa de uma maior necessidade de recursos humanos, financeiros e tecnológicos para os hospitais universitários, centrais e especializados.

A utilização da severidade do estado do doente, mediante o recurso ao Disease Staging, vem alterar substancialmente as conclusões atrás referidas.

De facto, embora se mantenha a relação directa entre a gravidade e os resultados dos hospitais, ou seja quanto maior a gravidade dos doentes maior a taxa de mortalidade e a demora média, os hospitais que tratam doentes mais graves são os de menor dimensão produtiva, afinal aqueles que na lógica

organizativa portuguesa correspondem maioritariamente a hospitais distritais de nível 1 (19 hospitais num total de 22).

Esta situação constitui desde já uma novidade no actual panorama hospitalar português, dado que se esperava que os hospitais com maior dimensão tratassem, à semelhança do que se verifica com a complexidade dos casos, igualmente os doentes mais graves.

Atendendo a que estes valores não parecem corresponder a nenhum fenómeno estrutural, visto que uma análise por doença vem comprovar uma associação directa e significativa entre a complexidade e a gravidade dos casos tratados, o comportamento distinto dos hospitais deve ser procurado noutras razões.

Para tal as principais hipóteses de investigação parecem apontar para questões associadas com a lógica organizacional dos hospitais portugueses, com a natureza específica das doenças e da respectiva gravidade dos casos, com uma eventual escolha de doentes por parte dos hospitais, ou ainda com a existência de diferentes práticas dos hospitais, em que nos de menor dimensão se concentram mais casos para tratamentos paliativos (Iezzoni et al, 1991b).

A consideração da primeira alternativa tem essencialmente em atenção aspectos estruturais da rede hospitalar portuguesa, na qual a lógica prevalecente poderá eventualmente residir numa concentração de recursos para se tratarem os casos mais complexos e os mais graves.

No capítulo da apresentação de Resultados foi referido que o nível de produção tem efeitos sobre a complexidade dos casos tratados, aumentando com incrementos da produção, sobre a gravidade dos casos tratados, diminuindo com incrementos da produção e sobre a diferença entre a complexidade e a gravidade dos casos tratados, igualmente aumentando com incrementos da produção.

Contudo, para além da análise desta variação média por níveis de produção, é igualmente interessante avaliar a estabilidade deste comportamento dentro de cada grupo produtivo.

Para tal e para os grupos de produção definidos e somente para todas as admissões serão apresentados os coeficientes de variação para a produção, para a gravidade dos casos tratados, para a complexidade dos casos tratados e para a respectiva diferença entre estes dois atributos (ver Quadro LVI).

A existência de razões estruturais e predominantemente localizadas em termos de oferta levará a uma heterogeneidade reduzida em cada grupo dos hospitais, essencialmente na complexidade dos casos tratados. A existência ou não de correspondência com a gravidade dos casos e com as diferenças entre estes dois atributos será igualmente analisada.

Quadro LVI
Variação na Produção Média, na Complexidade Média, na Gravidade Média e nas Diferenças entre Complexidade e Gravidade por Hospital por Quartis Todas as Admissões

	Produção Média	Gravidade Média	Complexidade Média	Diferença Complexidade Gravidade
Até 1º Quartil	0.32	0.72	1.52	0.59
Entre 1 e 2º Quartis	0.24	0.56	9.23	1.35
Entre 2 e 3º Quartis	0.14	0.41	8.28	0.98
Depois 3º Quartil	0.44	0.41	5.27	0.77

Estes elementos permitem as seguintes observações:

- Variação reduzida nos hospitais para a produção média, sendo a menor diferença encontrada no 2º grupo de hospitais com maior produção e maior diferença nos hospitais com maior produção;
- Variação igualmente reduzida na gravidade dos doentes tratados, embora com maior heterogeneidade nos hospitais com menor produção e com maior homogeneidade nos hospitais com maior produção;
- Grandes variações na complexidade dos casos tratados para todos os grupos de hospitais considerados, embora com menor expressão nos hospitais com menor produção;
- Grandes variações na diferença entre a complexidade e a gravidade dos casos tratados, embora com menor expressão nos hospitais de menor produção.

Daqui pode concluir-se que as características da oferta, nomeadamente no que se refere ao desenvolvimento tecnológico e à complexidade dos casos tratados não parece constituir um fenómeno estrutural visto que as maiores diferenças dentro de cada grupo de hospitais são exactamente encontradas nesta característica dos doentes tratados.

Deve ainda referir-se que atendendo à importância do assunto, no futuro sejam realizados outros estudos de investigação, tanto sobre o mecanismo da formação dos preços de pagamento em Portugal, como no aprofundamento desta análise, para casos médicos e cirúrgicos e por doença.

No entanto, a menor heterogeneidade encontrada, tanto para a gravidade dos casos tratados, como para a respectiva diferença com a complexidade, torna necessário que se passe a discutir as outras questões associadas com a natureza específica das doenças e com uma eventual escolha de doentes por parte dos hospitais.

Como foi referido no capítulo Resultados não parecem existir contradições entre a complexidade e a gravidade dos casos, tanto para o total das admissões, como para os casos médicos e cirúrgicos.

Atendendo a que estas indicações foram estabelecidas com base nas informações disponibilizadas pelos “z scores” construídos para a complexidade e para a gravidade estes podem estar a introduzir alguma perturbação na análise.

De facto, os “z scores” do Disease Staging podem não identificar a globalidade da gravidade do doente, tanto pelo facto de os estadios de gravidade poderem não ser equivalentes entre as diversas doenças ou porque a gravidade pode não estar a ser devidamente medida pelo tipo de escala (ordinal) utilizado.

Assim, interessa igualmente proceder a um novo apuramento para se avaliar a existência de potenciais contradições entre a complexidade e a gravidade dos casos por doença.

Para tal serão utilizados os Índices de Casemix dos DRGs e do Disease Staging, como “proxy” da complexidade dos casos e da gravidade dos casos, respectivamente.

De facto os “z scores” e os Índices de Casemix estão a disponibilizar informação ligeiramente diferente, tanto na complexidade, como na gravidade dos casos.

Para a complexidade e para o total das admissões o coeficiente de correlação entre estes dois indicadores é de -0.36. Para os casos cirúrgicos esta estatística é de -0.13, enquanto que nos casos médicos é de -0.35.

Por sua vez, para a gravidade dos casos tratados, os coeficientes de correlação são de -0.46 (significativo a 0.05), de -0.60 (significativo a 0.01) e de -0.35, respectivamente para todas as admissões, casos cirúrgicos e casos médicos.

No entanto, deve referir-se que para todas as associações, o sinal negativo na correlação, traduz o comportamento esperado, visto que os Índices de Casemix variam de uma forma directa com o aumento da gravidade ou da complexidade dos casos, enquanto que os “z scores” variam inversamente com o aumento da gravidade ou da complexidade dos casos.

Contudo, quando se pesquisam eventuais contradições entre a complexidade e a gravidade dos casos por doença, as conclusões mantêm-se inalteradas, ou seja quanto mais complexos, mais casos graves existem. Os coeficientes de correlação são de 0.79, de 0.81 e de 0.63 (todos significativos a 0.01), respectivamente para todas as admissões, casos cirúrgicos e casos médicos.

Assim, reforça-se o referido anteriormente, designadamente que os diferentes comportamentos dos hospitais em relação ao perfil das suas admissões, em

termos da complexidade e da gravidade dos casos tratados, não poderem ser justificados por razões estruturais, derivadas das especificidades das doenças.

Neste sentido, não se encontrando razões estruturais, tanto ligadas à estrutura produtiva dos hospitais, como associadas com as especificidades das doenças, para os diferentes comportamentos dos hospitais no que respeita ao seu perfil das admissões, pode concluir-se que este poderá traduzir uma política diferente dos mesmos no que se refere à escolha de doentes.

Deve igualmente ser evidenciado que esta análise global e genérica pode ser, pelo menos parcialmente infirmada quando se realizar uma avaliação mais desagregada, essencialmente por doença principal, pelo que no futuro se sugere que este tipo de estudos sejam realizados em Portugal, para que se possa caracterizar o comportamento dos hospitais, conhecer as suas razões justificativas e introduzir mecanismos para se corrigirem os problemas detectados.

Enquanto tal não for realizado, deve ter-se presente que este estudo indicia um comportamento distinto dos hospitais, no qual as organizações de saúde, com maior dimensão são as que simultaneamente tratam doentes mais complexos e menos graves.

Estas conclusões evidenciadas neste estudo, não eram até ao presente momento reconhecidas em Portugal, visto que se argumentava que os hospitais Centrais e Universitários (sendo aqueles que na generalidade apresentam maior capacidade de oferta e maior dimensão produtiva) por estarem no fim da linha da referência hospitalar, são os que tratam doentes mais graves (Serrão et al, 1998).

Este estudo permite evidenciar que a gravidade dos doentes tratados nos hospitais de maior dimensão, tem sido confundida por outros aspectos, designadamente a raridade dos casos tratados e ainda a sua complexidade.

No entanto, o estudo aponta ainda para o facto de os hospitais com doentes menos graves, menos complexos e menos graves e menos complexos apresentam melhores resultados brutos, quando os mesmos estão a ser medidos pela mortalidade e pela demora média.

Este estudo evidencia ainda que quanto maior é a diferença entre a complexidade e a gravidade dos casos tratados melhores são os resultados dos hospitais, expressos por taxas de mortalidade e demoras média mais baixas.

Esta situação é a esperada face a diversos estudos internacionais, onde são evidenciadas as vantagens de admitir doentes com menor gravidade e menor complexidade, não só ao nível dos resultados, mas igualmente sobre o desempenho financeiro dos hospitais (Hornbrook e Goldfarb, 1983; Gonnella,

Hornbrook e Louis, 1984; Schumaker et al, 1987; Thomas e Ascraft, 1989 e Barnum, Kutzin e Saxenian, 1995).

Todas estas questões suscitam, pelo menos, duas perspectivas suplementares de análise:

- Existe ou não uma relação entre o perfil das admissões dos hospitais e os seus resultados, depois de ajustados pelo risco?
- Quais os mecanismos que devem ser utilizados para se corrigir ou melhorar os problemas diagnosticados?

Assim, à semelhança do realizado para os resultados brutos, a mortalidade e a demora média, será feita uma análise para os resultados dos hospitais depois de ajustados pelo risco.

As variáveis para caracterizar o perfil das admissões dos hospitais são as mesmas:

- Gravidade dos doentes – Índice de Casemix do Disease Staging;
- Complexidade dos doentes – Índice de Casemix dos DRGs;
- Gravidade e Complexidade dos doentes – Índice de Casemix Composto;
- Diferença entre a Gravidade e a Complexidade – Índice de Casemix do Disease Staging – Índice de Casemix dos DRGs.

Os indicadores de resultados ajustados pelo risco são os seguintes:

- Efectividade – Comparação entre a mortalidade observada e a mortalidade prevista (“ z score”);
- Eficiência – Comparação entre a demora média observada e a demora média prevista (“ z score”);
- Global – Efectividade e Eficiência com a mesma ponderação (média dos (“ z scores”).

No Quadro LVII são apresentados os coeficientes de correlação entre o perfil das admissões dos hospitais e os respectivos resultados ajustados pelo risco.

Quadro LVII
Coeficientes de Correlação por Hospital entre
Admissões e Resultados Ajustados pelo Risco

	Efectividade	Eficiência	Global
Complexidade	-0.060	0.031	-0.039
Gravidade	0.109	0.015	0.102
Composto	0.061	0.025	0.064
Diferença	0.148	-0.003	0.128

Como primeira observação deve referir-se que nenhuma das correlações é significativa, pelo que não existe nenhuma associação entre a gravidade, a complexidade, a gravidade e complexidade e a diferença entre estes dois atributos e o desempenho dos hospitais, tanto para a efectividade, como para a eficiência e naturalmente para os resultados globais.

Esta situação é completamente diferente da retratada quando se consideram somente os resultados brutos, visto que genericamente os hospitais apresentavam melhor desempenho quando tratavam doentes menos graves e complexos ou maximizavam a diferença entre estes dois atributos.

Estas conclusões podem ser consideradas como esperadas, atendendo a que os hospitais estavam classificados de uma forma distinta quando se consideravam os resultados brutos e os resultados ajustados pelo risco.

Nesta conformidade, pode referir-se que a introdução do ajustamento pelo risco, elimina pelo menos um dos problemas decorrentes de diferentes políticas de admissão por parte dos hospitais – a apresentação de melhores indicadores e de resultados da sua actividade.

No entanto, subsistem ainda alguns problemas que devem ser discutidos, numa tentativa de se perspectivarem soluções para o futuro. Estas questões estão intimamente associadas com a política de saúde, designadamente no que se refere à universalidade do sistema, à acessibilidade às organizações de saúde e com o financiamento de hospitais, bem como com aspectos de administração em saúde, designadamente com o planeamento e gestão dos recursos humanos e tecnológicos.

Atendendo à importância que envolvem estes tópicos serão discutidos numa fase posterior deste estudo – Implicações para a reorganização do sector da saúde.

6.2.2. Desempenho

6.2.2.1. Efectividade

No capítulo Resultados foram identificados alguns aspectos de entre os quais se destacam:

- Existem indicações completamente distintas quando se analisa o desempenho dos hospitais (efectividade) em função dos seus resultados brutos e ajustados pelo risco. Estas conclusões são válidas para qualquer das perspectivas de análise, ou seja, todas as admissões, casos cirúrgicos e médicos e por doença;
- Quando se pretende ajustar pelo risco existem, mais uma vez, indicações distintas entre os métodos directo e indirecto para a sua concretização, sendo estas conclusões igualmente válidas para todas as perspectivas de análise;
- Observou-se ainda um comportamento bastante heterogéneo dos hospitais quando a avaliação do desempenho é feita de uma forma mais desagregada. Ou seja, os hospitais em termos genéricos apresentam comportamentos distintos entre os casos cirúrgicos e médicos e, de uma forma mais nítida, por doença.

Estes aspectos merecem algumas reflexões suplementares.

Em primeiro lugar, deve referir-se que os argumentos apresentados pelos hospitais que os seus resultados brutos não devem ser comparados, essencialmente devido às características dos seus doentes, onde avulta a gravidade, parecem ter cabimento.

Na realidade, alguns hospitais apresentam taxas de mortalidade bastante elevadas, simplesmente pelo facto da média da gravidade dos doentes tratados ser mais elevada. Bons exemplos ilustrativos desta situação são os hospitais 65, 15, 7, 4, 11, 6, 73, 79, 13 e 10.

Na situação contrária, isto é organizações de saúde cujos resultados brutos são bastante melhores que os ajustados pelo risco, encontram-se os hospitais 29, 18, 28, 20, 85, 31, 47, 56, 55 e 82. Assim, estes hospitais aparecem incorrectamente identificados com bom desempenho, sendo as suas baixas taxas de mortalidade devidas à menor gravidade dos doentes tratados.

Outra forma de ilustrar este comportamento deriva da comparação entre os 10 hospitais com taxas de mortalidade mais elevadas (resultados brutos), com os seus resultados ajustados pelo risco, visto que nenhum deles ocupa as 10 últimas posições na sua efectividade. Quando se analisam os 10 hospitais com taxas de mortalidade mais baixas observa-se que somente o hospital 41 aparece

igualmente entre os 10 hospitais com melhor desempenho após ajustamento pelo risco.

Em segundo lugar, deve referir-se que os métodos directo e indirecto para se ajustarem os resultados pelo risco conduzem igualmente a indicações diferentes sobre o desempenho dos hospitais.

Isto implica que, quando se estabelecerem critérios para se avaliar o desempenho dos hospitais, deve ser indicada qual a metodologia utilizada, por mais que não seja, pelo facto de se identificarem hospitais distintos com melhor e pior desempenho em função das respectivas metodologias.

Os hospitais 10, 78, 13, 81 e 68 são os que aparecem mais beneficiados pelo ajustamento directo ("z score"), enquanto que os hospitais 18, 20, 21, 22 e 17 são os mais prejudicados por este tipo de ajustamento.

Mesmo assim, os dois métodos concordam em 5 hospitais entre os 10 com pior desempenho, mas somente em 1 para os 10 hospitais com melhor desempenho.

Conforme foi referido existem argumentos teóricos que justificam a utilização do método directo como a melhor metodologia para se compararem os valores observados com os esperados, essencialmente porque toma em consideração o diferencial da gravidade dos doentes tratados por cada hospital (Shwartz, Ash e Iezzoni, 1997).

Para os hospitais portugueses esta situação é bem patente, visto que para os 10 melhores hospitais quando se utiliza o método indirecto a gravidade em relação à média de todos os hospitais é de 0.45, enquanto que para os 10 hospitais com pior desempenho este ratio é de 1.25.

Com o método directo, estes ratios são de 1.01 e de 1.10, respectivamente para os 10 melhores e piores hospitais em termos de desempenho.

Embora se observe que os hospitais com pior desempenho continuem a apresentar um resultado idêntico, quando se utiliza o método directo, o diferencial encontrado na gravidade entre os dois grupos de hospitais, é só por si suficientemente esclarecedor quanto ao facto de o método directo privilegiar os hospitais que tratam doentes com maior gravidade.

Por outro lado, a utilização do método indirecto, conduz a que para os melhores hospitais 5 são do Tipo II (num máximo de 6), 2 dos Tipos III e IV e 1 do Tipo V, com uma dimensão média da produção de 12720 doentes.

Com o método directo encontram-se 5 hospitais do Tipo I, 4 do Tipo III e 1 do Tipo V (o único comum às duas metodologias), mas agora com uma dimensão

média da produção de 59169 doentes (somente a título indicativo a produção média por hospital é de 26309 doentes).

Como foi referido no capítulo Resultados estes dois métodos têm implicações distintas na identificação dos hospitais com melhor desempenho, podendo inclusivamente afirmar-se que o método directo contribui para uma melhor comparação entre hospitais, dispensando a necessidade de se estabelecerem metodologias para garantir a uniformidade dos hospitais baseada em critérios de oferta.

Recorde-se a este propósito que esta é a metodologia utilizada em Portugal até ao momento, a qual para além de disponibilizar informações erróneas sobre a actividade do hospital, contribui ainda mais para se internalizar a avaliação do desempenho dos hospitais.

Finalmente, deve evidenciar-se o desempenho distinto dos hospitais, tanto entre os casos cirúrgicos e médicos, como por doença. A este propósito foi inclusivamente referido anteriormente que existem hospitais cujas diferenças no desempenho por doença são por vezes superiores a diferenças encontradas entre diferentes hospitais.

No que se refere à diferença entre os casos médicos e cirúrgicos podem fazer-se as seguintes observações:

- Os hospitais 87, 88, 25, 28 e 67 são os que apresentam maiores discrepâncias no desempenho entre os casos cirúrgicos e os médicos, com melhor desempenho nos primeiros;
- Com melhor desempenho nos casos médicos do que nos cirúrgicos, as maiores discrepâncias são encontradas nos hospitais 13, 44, 7, 60 e 24;
- Por sua vez, os hospitais 9, 18, 24, 61, 68, 33, 42 e 43 são os que apresentam um desempenho idêntico entre os casos médicos e os cirúrgicos.

Por outro lado, os 5 hospitais que comparativamente apresentam melhor desempenho nos casos cirúrgicos do que nos médicos, têm igualmente um fraco desempenho global. Entre estes o hospital 25 é o que aparece mais bem classificado para todas as admissões, ocupando a 60^a posição.

No entanto, para os 5 hospitais que comparativamente apresentam melhor desempenho nos casos médicos, a situação é completamente distinta, visto que o hospital 13 ocupa a 2^a posição para todas as admissões, enquanto que o hospital 60 ocupa a 46^a posição.

Mais uma vez, para os 8 hospitais que apresentam desempenhos semelhantes entre os casos médicos e cirúrgicos não se vislumbra nenhum padrão. Enquanto

que os hospitais 9 e 61 têm bom desempenho para todas as admissões, os hospitais 33 e 43 têm mau desempenho para todas as admissões e o hospital 42 apresenta um desempenho intermédio.

Para além destes aspectos existem outros que merecem uma discussão suplementar, designadamente os relacionados com algumas questões teóricas e com potenciais contradições encontradas neste estudo.

Para as primeiras, questões conceptuais irão ser analisados alguns tópicos referenciados na literatura internacional, como relevantes para diferentes desempenhos dos hospitais. Dentro destes destacam-se o tipo de hospital, a localização do hospital e a dimensão do hospital (Blumberg, 1986; Iezzoni et al, 1997; Birkmeyer et al, 2002).

Para as questões relacionadas com os resultados deste estudo, avultam as relativas a eventuais desempenhos distintos entre os hospitais, tanto para casos cirúrgicos e médicos, como para diferenças entre as doenças.

Em relação às questões conceptuais serão feitas análises de regressão em que as variáveis dependentes são os resultados brutos (taxa de mortalidade) e os resultados ajustados pelo risco, pelo método directo (“z score”).

As variáveis explicativas são as seguintes:

- Dimensão – medida pelo volume total da produção por hospital (doentes saídos). Para se pesquisarem eventuais relações não lineares, será igualmente utilizada produção elevada ao quadrado;
- Tipo de hospital – será utilizada uma variável “dummy”, identificando os seguintes tipos de hospital: Central e Universitário; Especializado; Distrital; Distrital de Nível 1 e Institutos Portugueses de Oncologia (IPOs).
- Localização do hospital – será utilizada uma variável “dummy”, identificando os hospitais em função da respectiva Região de Saúde: Norte, Centro, Lisboa e Vale do Tejo, Alentejo e Algarve. A este respeito, embora a distinção entre a localização urbana e rural também seja relevante, optou-se por esta operacionalização, para existir uma maior correspondência com a lógica de gestão em vigor em Portugal.

Para estas análises de regressão serão analisados o total das admissões e os casos cirúrgicos e médicos.

Os principais resultados são os seguintes:

Total das Admissões

Resultados brutos:

R^2 (ajustado) = 0.471; Estatística F = 8.738 (significativa a 0.05).

Figura 59
Resultados do Modelo: Mortalidade; Total das Admissões

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Sim
Produção (termo ao quadrado)	Não
Centrais e Universitários	Não
Especializados	Sim
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Sim
Norte	Sim

Os resultados brutos dos hospitais estão a ser explicados pela produção, aumentando a taxa de mortalidade com a diminuição da produção, pelo Tipo de hospital, embora só seja significativo o coeficiente de regressão para os hospitais especializados e pela localização do hospital, sendo só significativos os coeficientes de regressão para as Regiões Centro e Norte. Deve ainda referir-se que não existe multicolinearidade no modelo, visto que o “Eigenvalue” é de 0.18, com um Índice de Condição de 13.794 e que os resíduos são normais com a estatística do Teste de Kolmogorov-Smirnov a ser igual a 0.863 (probabilidade de significância > 0.05).

Estes resultados parecem estar de acordo com o padrão esperado, ou seja a taxa de mortalidade é afectada pela dimensão dos hospitais, variando de uma forma inversa e pelo tipo de hospital, bem como pela sua localização, implicando resultados diferentes para os hospitais.

Analisando agora os resultados ajustados pelo risco os valores são os seguintes:

R^2 (ajustado) = 0.371; Estatística F = 6.131 (significativa a 0.05).

Figura 60
Resultados do Modelo: Efectividade; Total das Admissões

Variáveis explicativas	Significância (< 0.05)
Constante	Não
Produção (termo simples)	Não
Produção (termo ao quadrado)	Sim
Centrais e Universitários	Não
Especializados	Não
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Sim
Norte	Sim

Os resultados ajustados pelo risco apenas estão a ser explicados pela produção e pela localização dos hospitais, sendo neste particular somente significativos os coeficientes de regressão para as Regiões Centro e Norte. Mais uma vez não existe multicolinearidade no modelo, sendo os valores do “Eigenvalue” e do Índice de Condição, iguais aos do modelo anterior. Os resíduos são normais, assumindo a estatística do Teste de Kolgomorov-Smirnov um valor de 1.126 (probabilidade de significância > 0.05).

No entanto, os resultados são diferentes dos observados para os resultados brutos, designadamente com o desaparecimento da influência do Tipo de hospital e na relação entre os resultados e dimensão dos hospitais.

Para o primeiro aspecto, estes resultados práticos confirmam a importância do ajustamento pelo risco para controlar as características dos doentes e desta forma para se proceder a uma análise mais criteriosa do desempenho hospitalar.

Mais uma vez, se reforça o anteriormente referido, no qual a comparação dos hospitais baseada em critérios de oferta não é a forma mais correcta e pode inclusivamente introduzir alguns enviesamentos na avaliação do desempenho dos hospitais.

Por outro lado, com o ajustamento pelo risco, os hospitais são mais efectivos com maiores e menores dimensões produtivas, o que parece contrariar o esperado na literatura internacional, nomeadamente uma relação directa entre quantidade produzida e resultados de saúde.

Tal facto, para além de necessitar de análises mais aprofundadas no futuro, suscita igualmente as reservas sobre a ausência de racionalidade e de gestão nos hospitais portugueses.

Finalmente, deve ter-se em conta que os hospitais que face à sua dimensão produtiva e localização apresentam um melhor desempenho em termos de resultados ajustados pelo risco são os números 15, 61, 13, 7, 48, 5, 86, 4, 2 e 53, sendo 6 do Tipo I, 3 do Tipo IV e 1 do Tipo V. Estes hospitais estão localizados na Região IV (5), na Região III (2) e nas Regiões II e V (1).

É igualmente interessante referir que estes 10 hospitais ocupam as 25 primeiras posições na efectividade. No entanto o hospital 9, o mais efectivo, ocupa somente a 44^a posição quando se compara este valor com o esperado face à sua localização e dimensão produtiva. A situação do hospital 10 merece igualmente relevo, visto que apesar de ocupar a 4^a posição na efectividade, está entre os 20 piores face aos valores esperados.

Face à sua dimensão produtiva e localização os hospitais 12, 1, 59, 43, 88, 62, 55, 30, 51 e 40 são os que apresentam pior desempenho. Destes hospitais 7 são

do Tipo III, 2 do Tipo I e 1 do Tipo V, estando 4 localizados na Região de Saúde IV, 2 nas Regiões III e V e 1 nas Regiões I e II.

Visto que estes hospitais, ocupam igualmente as últimas posições na ordenação da efectividade, não se encontram grandes discrepâncias entre a efectividade e o desempenho esperado face à sua produção e localização.

Casos Cirúrgicos

Resultados brutos:

R^2 (ajustado) = 0.708; Estatística F = 21.606 (significativa a 0.05).

Figura 61
Resultados do Modelo: Mortalidade; Casos Cirúrgicos

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Não
Produção (termo ao quadrado)	Não
Centrais e Universitários	Sim
Especializados	Sim
Distritais de Nível 1	Sim
IPOs	Sim
Alentejo	Não
Algarve	Não
Centro	Sim
Norte	Sim

Os resultados brutos estão somente a ser explicados pelo Tipo de hospital e pela sua localização. Mais uma vez não existe multicolinearidade no modelo e os resíduos são normais.

Para os resultados ajustados pelo risco os valores são os seguintes:

R^2 (ajustado) = 0.158; Estatística F = 2.595 (significativa a 0.05).

Não existe multicolinearidade no modelo e os resíduos são normais, mas a efectividade para os casos cirúrgicos é somente explicada pelo nível produtivo e pelo Tipo de hospitais (somente os IPOs).

Para os casos cirúrgicos os resultados da análise de regressão são ainda mais surpreendentes.

Em primeiro lugar, o poder explicativo do modelo é bastante reduzido, sendo o R^2 ajustado somente de 0.158.

Em segundo lugar, ao contrário do que sucedeu para os resultados brutos, encontra-se um efeito da dimensão produtiva, mais uma vez numa forma de U invertido, embora a maior efectividade seja encontrada nos hospitais com maior produção.

Em terceiro lugar, ao contrário do que se verificou para o total das admissões não existe importância da localização dos hospitais, verificando-se contudo que os IPOs são relevantes para o diferente desempenho dos hospitais.

Figura 62
Resultados do Modelo: Efectividade; Casos Cirúrgicos

Variáveis explicativas	Significância (< 0.05)
Constante	Não
Produção (termo simples)	Não
Produção (termo ao quadrado)	Sim
Centrais e Universitários	Não
Especializados	Não
Distritais de Nível 1	Não
IPOs	Sim
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Não

Os hospitais 48, 61, 6, 5, 4, 25, 39, 41, 2 e 15 são os que apresentam efectividade superior à esperada em função da sua produção e tipo de hospital. Estes hospitais são 6 do Tipo I e 4 do Tipo III, estando 5 localizados na Região de Saúde IV, 3 na Região III e 2 na Região V.

Para os 10 hospitais com valores melhores que o esperado verifica-se que já ocupavam as 20 primeiras posições em termos de efectividade.

Os hospitais 1, 12, 40, 13, 36, 43, 60, 16, 33 e 44 são os que apresentam pior efectividade face à esperada em função da sua produção e tipo de hospital. Encontram-se 6 hospitais do Tipo III e 4 do Tipo I, estando 4 localizados na Região de Saúde IV, 3 na Região V, 2 na Região 3 e o restante na Região II.

Observa-se igualmente uma correspondência entre os níveis de efectividade e os valores esperados em função da produção e do tipo de hospital.

Casos Médicos

Resultados brutos:

R^2 (ajustado) = 0.565; Estatística F = 12.310 (significativa a 0.05).

Figura 63
Resultados do Modelo: Mortalidade; Casos Médicos

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Não
Produção (termo ao quadrado)	Não
Centrais e Universitários	Sim
Especializados	Sim
Distritais de Nível 1	Sim
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Sim
Norte	Sim

O modelo não apresenta multicolinearidade e os resíduos são normais, sendo os resultados brutos exclusivamente explicados pelo Tipo de hospital (somente os Centrais, Especializados e Distritais de nível 1 apresentam coeficientes de regressão significativos) e pela localização do hospital (somente nas Regiões Centro e Norte os coeficientes de regressão são significativos).

Para os resultados ajustados pelo risco os valores são os seguintes:

R^2 (ajustado) = 0.417; Estatística F = 7.228 (significativa a 0.05).

Figura 64
Resultados do Modelo: Efectividade; Casos Médicos

Variáveis explicativas	Significância (< 0.05)
Constante	Não
Produção (termo simples)	Não
Produção (termo ao quadrado)	Sim
Centrais e Universitários	Não
Especializados	Não
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Sim
Norte	Sim

Mais uma vez, para os casos médicos, os resultados ajustados pelo risco diferem dos resultados brutos, passando a produção a ter importância (a efectividade apresenta igualmente uma forma de U invertido, embora com maior expressão nos hospitais de maior dimensão) e considerando ainda a localização.

Os hospitais 15, 13, 61, 7, 86, 5, 54, 32, 53 e 4 são os que apresentam melhores valores face à efectividade esperada em função da sua produção e localização. São 5 hospitais do Tipo I, 4 do Tipo III e 1 do Tipo V, estando 5 localizados na Região de Saúde IV, 3 na Região III e 1 nas Regiões II e V.

Na situação contrária, efectividade pior que os valores esperados, encontram-se os hospitais 12, 59, 43, 88, 62, 55, 30, 1, 10 e 26. Estes hospitais estão localizados na Região de Saúde IV (5), na Região V (3) e nas Regiões I e III (1), sendo 6 do Tipo III, 3 do Tipo e I do Tipo V.

Em síntese pode afirmar-se:

- Existe um efeito dimensão da produção para os resultados ajustados pelo risco, tanto para o total das admissões, como para os casos cirúrgicos e médicos;
- O ajustamento que melhor traduz esta relação assume uma forma quadrática, numa forma de U invertido, pelo que a melhor efectividade é encontrada nos hospitais com maiores e menores volumes de produção, embora com maior intensidade para a maior produção. Esta afirmação é igualmente válida para o total das admissões e para os casos cirúrgicos e médicos;
- A localização dos hospitais, medida pela Região de Saúde, tem igualmente influência para o total das admissões e para os casos médicos;
- Para os casos cirúrgicos esta dimensão não está presente, sendo no entanto relevante o tipo de hospital.

Esta análise macro, embora identifique a relevância da produção para a obtenção de níveis de efectividade não é totalmente conclusiva para efeitos de planeamento e da administração das organizações de saúde, visto que os melhores valores podem ser obtidos com altos e baixos níveis de produção.

No entanto, apesar destas conclusões, parece igualmente evidente que os hospitais com maior dimensão produtiva são mais predominantes entre aqueles que apresentam maior efectividade, pelo que se torna importante a realização de estudos mais desagregados, essencialmente ao nível da doença, para que desta forma se possa perspectivar o desenho futuro dos serviços e dos hospitais.

Por outro lado, torna-se igualmente importante equacionar este tipo de análise para a eficiência e para o desempenho global dos hospitais, para que se possam detectar eventuais contradições e tornar mais explícita a dimensão óptima dos hospitais.

No capítulo apresentação de resultados foi igualmente identificado um diferente desempenho dos hospitais para os casos cirúrgicos e médicos, bem expresso pelo *K de Cohen* (0.035). Ou seja, os hospitais são classificados de forma completamente diferente em função do tipo de casos (cirúrgicos ou médicos) que estão a tratar.

Neste sentido, as questões mais relevantes a discutir são:

- Será que este fenómeno é estrutural?
- Existe ou não um padrão no comportamento dos hospitais que justifique o diferente desempenho para casos cirúrgicos e médicos?

Em relação ao primeiro aspecto, para a pesquisa sobre o facto de o fenómeno ser ou não estrutural serão feitas as seguintes análises:

- Correlação entre a efectividade dos casos cirúrgicos e médicos, através da padronização directa (“z scores”);
- Correlação entre as ordenações dos hospitais para os casos cirúrgicos e médicos, através da padronização directa;
- Matriz da ordenação dos hospitais por efectividade, para os casos cirúrgicos e médicos.

O coeficiente de correlação entre a efectividade dos hospitais para os casos cirúrgicos e médicos é de 0.489 (significativo a 0.01), pelo que se pode concluir que a efectividade dos casos cirúrgicos aumenta à medida que aumenta a dos casos médicos e vice-versa.

O coeficiente de correlação de Spearman para a ordenação dos hospitais por nível de efectividade para os casos médicos e cirúrgicos é de 0.405 (significativo a 0.01), o que mais uma vez parece traduzir uma associação directa na classificação dos hospitais para os casos médicos e cirúrgicos.

Neste sentido, embora os hospitais sejam classificados de forma diferente segundo o seu desempenho para os casos cirúrgicos e médicos, tudo parece indicar que este fenómeno não é estrutural, visto que, na generalidade, os hospitais que são mais efectivos nos casos cirúrgicos, também o são nos casos médicos.

No Quadro LVIII é apresentada a matriz da ordenação dos hospitais por efectividade para os casos cirúrgicos e médicos por quartis.

Assinalados a azul estão os 38 hospitais que estão classificados no mesmo grupo de desempenho. A vermelho estão identificados 18 hospitais com maiores discrepâncias no desempenho entre os casos cirúrgicos e os médicos. Para os restantes 30 hospitais considera-se que as diferenças não são muito relevantes.

Esta análise vem confirmar as indicações retiradas da análise de correlação, visto que 38 hospitais (44.19%) apresentam o mesmo desempenho para os casos cirúrgicos e médicos, enquanto que somente 20.96% (18 hospitais) apresentam desempenhos completamente distintos para estes dois grandes conjuntos de doentes.

Quadro LVIII
Efectividade
Ordenação do Desempenho dos Hospitais, por Quartis
Casos Cirúrgicos e Casos Médicos

Médicos				
Cirúrgicos	Até 1º Quartil	Entre o 1º e o 2º Quartis	Entre o 2º e o 3º Quartis	Depois do 3º Quartil
Até 1º Quartil	11	7	2	2
Entre o 1º e o 2º Quartis	5	7	7	2
Entre o 2º e o 3º Quartis	2	1	10	8
Depois do 3º Quartil	4	6	2	10

Para se analisar uma eventual existência de um padrão para o diferente comportamento dos hospitais para os casos médicos e cirúrgicos, serão utilizados os três grupos de hospitais: os hospitais com menores discrepâncias entre os casos cirúrgicos e médicos, os hospitais com discrepâncias médias nestes mesmos casos e os hospitais com maiores discrepâncias.

Estes grupos serão testados em função dos seguintes aspectos:

- Desempenho dos hospitais;
- Tipo do hospital;
- Localização do hospital;
- Produção do hospital;
- “Mix” de casos tratados.

Em relação ao desempenho dos hospitais (ver elementos do Quadro LVIII) observa-se que para os 38 hospitais que apresentam o mesmo comportamento nos casos cirúrgicos e médicos, existe uma distribuição quase idêntica por níveis de desempenho, visto que 11 se encontram entre os melhores hospitais, 10 entre os piores e 7 e 10 para cada um dos níveis intermédios de desempenho (efectividade).

No entanto, para os 18 hospitais com maiores discrepâncias no desempenho entre os casos médicos e cirúrgicos a situação é diferente. De facto, para os 22 hospitais com pior desempenho nos casos cirúrgicos, 4 encontram-se entre os melhores nos casos médicos e 6 no 2º grupo com melhor efectividade igualmente para estes casos. Para estes 10 hospitais constata-se que 8 são do Tipo III e 2 do Tipo I, estando 4 localizados na Região de Saúde V, 3 na Região III, 2 na Região IV e 1 na Região I.

Estes valores parecem indiciar que nos hospitais do Tipo III e na Região V e eventualmente na Região III a existência de comportamentos diferentes dos hospitais para os casos cirúrgicos e médicos, os quais, no mínimo, são distintos do padrão nacional.

Para melhor se caracterizar este fenómeno, sugere-se que no futuro sejam realizados diversos estudos, tendentes a investigar se estes comportamentos se devem ou não a diferentes níveis de efectividade entre os casos médicos e cirúrgicos.

Neste momento e atendendo somente ao desempenho global dos hospitais, pode afirmar-se que parece não existir um padrão justificativo para o diferente desempenho dos hospitais para os casos cirúrgicos e médicos.

No Quadro LIX são apresentados os hospitais segundo a heterogeneidade do desempenho entre os casos cirúrgicos e médicos por tipo de hospital.

Quadro LIX
Efectividade
Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos
por Tipo de Hospital

	Tipo I	Tipo II	Tipo III	Tipo IV	Tipo V
Pequenas Variações	8 (7)	3 (3)	18 (18)	8 (9)	1 (1)
Variações Médias	3 (5)	3 (2)	11 (14)	13 (8)	0 (1)
Grandes Variações	4 (3)	0 (1)	11 (8)	1 (5)	2 (1)

Entre parênteses encontram-se os valores esperados se a distribuição fosse equitativa por todas as células, observando-se o seguinte:

- Para os hospitais do Tipo I existem menos 2 hospitais que o esperado no grupo das variações médias e mais 1 hospital que o esperado para aqueles que apresentam comportamento menos e mais homogéneo;
- Para os hospitais do Tipo II existe menos 1 hospital que o esperado no grupo com grandes variações e mais que um que o esperado no grupo de variações médias;
- Para os hospitais do Tipo III existem mais 3 hospitais que o esperado para o grupo de grandes variações e menos 3 que o esperado para o grupo de variações médias;
- Nos hospitais do Tipo IV existem mais 5 hospitais que o esperado no grupo de variações médias e menos 4 e 1 que o esperado, respectivamente para os grupos de grande variação e de variação reduzida;

- Nos hospitais do Tipo V existe 1 hospital a mais do que o esperado no grupo de grandes variações e menos 1 do que o esperado para a variação média.

Estes resultados, apesar das diferenças encontradas nos hospitais do Tipo III e do Tipo IV, parecem igualmente confirmar que não existem diferenças substanciais por Tipo de hospital para o diferente desempenho entre os casos cirúrgicos e médicos, pelo que mais uma vez não se descortina nenhum padrão justificativo de tal comportamento.

No Quadro LX são apresentados os hospitais segundo a heterogeneidade do desempenho entre os casos cirúrgicos e médicos por Região de Saúde.

Quadro LX
Efectividade
Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos
por Região de Saúde

	Região I	Região II	Região III	Região IV	Região V
Pequenas Variações	2 (2)	1 (1)	13 (12)	13 (12)	9 (11)
Variações Médias	2 (2)	1 (1)	8 (9)	9 (9)	10 (9)
Grandes Variações	1 (1)	1 (1)	5 (5)	5 (6)	6 (5)

Entre parênteses encontram-se os valores esperados se a distribuição fosse equitativa por todas as células.

Atendendo a que somente existem algumas discrepâncias nas Regiões de Saúde III, IV e V e em que somente nesta última Região existem menos 2 hospitais que o esperado no grupo de hospitais de variação reduzida, tudo parece indiciar que não existe nenhum padrão regional justificativo do diferente desempenho dos hospitais para os casos cirúrgicos e médicos.

No entanto, como foi referido anteriormente, recomenda-se a realização futura de estudos mais detalhados para melhor caracterizar este fenómeno.

No Quadro LXI são apresentadas a produção média e respectivo coeficiente de variação, e a percentagem de casos cirúrgicos, para todos os hospitais, bem como para os grupos de variação entre casos cirúrgicos e médicos anteriormente definidos.

No que se refere à produção dos hospitais, podem evidenciar-se alguns aspectos distintivos nos diversos grupos de hospitais em função do desempenho distinto para casos médicos e cirúrgicos.

Quadro LXI
Produção Média e Coeficiente de Variação; % de Casos Cirúrgicos
Todos os Hospitais e por Grupos de Variação

	Produção	Coeficiente de Variação	% Casos Cirúrgicos
Hospitais	26841	0.88	38.9
Pequenas Variações	30399	0.90	39.5
Variações Médias	17326	0.96	42.0
Grandes Variações	35189	0.61	35.2

Assim, nos hospitais com maior heterogeneidade observa-se uma dimensão produtiva bastante superior do que a média dos hospitais. A heterogeneidade produtiva neste grupo de hospitais, medida pelo coeficiente de variação é a mais reduzida de todas as perspectivas em análise.

No grupo de variações médias na efectividade entre os casos cirúrgicos e médicos, observa-se a produção média mais baixa, embora com a maior heterogeneidade dos hospitais dentro do grupo.

No grupo com maior homogeneidade na efectividade entre os casos cirúrgicos e médicos observa-se igualmente uma produção média superior à da totalidade dos hospitais, igualmente com uma variação entre os hospitais do grupo bastante elevada.

A análise de variância simples (“one-way anova”) permite afirmar que as diferenças na produção média dos grupos de hospitais são significativas. Esta tendência mantém-se quando se analisam somente os grupos de variação média e grande e de variação pequena e média. Contudo, tal não se observa (diferença entre médias da produção) quando se comparam somente os grupos de variação pequena e grande.

Estes elementos permitem concluir pela influência da produção dos hospitais no desempenho distinto dos hospitais entre casos médicos e cirúrgicos. No entanto, o facto de esta não estar presente entre os dois grupos mais distintos neste comportamento (pequenas e grandes variações), suscita a interrogação se esta mesma influência poderá não ser uma razão explicativa, mas sim uma consequência de outras características dos hospitais.

Um dos factores que poderá contribuir para estas diferenças poderá ser a proporção de casos cirúrgicos tratados por hospital. No Quadro LXI observa-se que existem menos casos cirúrgicos tratados no grupo de hospitais com maior heterogeneidade na efectividade entre estes e os casos médicos e que a proporção de maiores casos cirúrgicos é encontrada nos hospitais com variações médias na efectividade.

Por outro lado, a análise da variância introduz novos elementos, visto que globalmente permite concluir que não existem diferenças na percentagem de casos cirúrgicos tratados entre os diversos grupos de hospitais. Esta situação mantém-se entre os grupos de média e grande variação e de média e pequena variação. Mas entre os grupos de hospitais de pequena e grande variação existem diferenças significativas na proporção de casos cirúrgicos tratados.

Neste sentido, tudo parece apontar que a maior dimensão produtiva dos hospitais pode conduzir a diferentes efectividades para casos cirúrgicos e médicos, embora a proporção de casos cirúrgicos tratados por hospital também parece constituir um factor relevante para este comportamento.

Mais uma vez, se afirma a necessidade de se realizarem estudos mais aprofundados para melhor se esclarecer o diferente desempenho dos hospitais na efectividade dos cuidados prestados, tendo-se igualmente presente que muitos destes resultados podem igualmente resultar de diferenças na gestão destas organizações de saúde.

No capítulo Resultados foram evidenciados alguns aspectos, quando se analisa a efectividade dos hospitais por doença, nomeadamente a existência de distintos hospitais com melhor efectividade por doença, tanto para a 1ª posição, como para as cinco primeiras posições ou ainda para os hospitais com pior efectividade (igualmente para a última posição, ou para as cinco últimas posições) e uma grande heterogeneidade na efectividade dentro de cada hospital em função das diversas doenças.

Estas questões conduzem à necessidade de reforço da gestão e dos mecanismos de regulação dos hospitais e suscitam novos problemas para o financiamento das organizações de saúde e para o exercício da liberdade de escolha por parte dos consumidores.

Para melhor se visualizarem estas dimensões serão apresentados para os 10 hospitais com melhor efectividade global, para os 10 com pior efectividade global, bem como para as organizações com maior heterogeneidade na efectividade, o resumo do seu desempenho para os casos cirúrgicos e médicos e por doença.

No Quadro LXII são identificados os desempenhos dos 10 melhores hospitais para todas as admissões, nos casos cirúrgicos e médicos e nas doenças, categorizadas pelo Grande Agrupamento de Doenças do Disease Staging.

Na primeira coluna, entre parênteses está identificado o número de hospitais que foram classificados em cada perspectiva de análise. A este propósito, recorde-se que não são considerados os hospitais que trataram menos de 100 doentes (situações que são identificadas por "na").

Quadro LXII
Efectividade Global, 10 Melhores Hospitais
Desempenho para os Casos Cirúrgicos e Médicos e por Doença
Ordenação

	H9	H13	H61	H10	H2	H15	H48	H32	H41	H86
Total (88)	1	2	3	4	5	6	7	8	9	10
Cirúrgicos (86)	1	82	3	9	8	18	2	19	6	5
Médicos (88)	1	2	4	3	5	5	12	9	10	13
BR (56)	4	1	16	7	3	45	38	44	15	2
CN (84)	62	2	4	1	7	10	23	6	21	47
CV (77)	1	62	10	2	3	15	14	11	9	na
DE (82)	72	81	3	46	47	39	23	12	9	2
EN (63)	1	3	20	6	4	37	5	22	10	2
GI (83)	1	2	5	3	11	4	6	13	14	7
GY (69)	2	60	6	15	1	14	9	41	21	3
HB (80)	1	2	3	12	63	13	4	14	9	11
HE (57)	2	14	28	38	16	4	35	na	17	40
IM (24)	1	8	na	22	3	2	20	na	na	na
LY (21)	1	7	8	2	6	na	na	na	na	19
MG (71)	12	3	16	9	11	55	4	47	28	45
MS (82)	21	80	1	10	7	43	4	24	17	36
ND (82)	1	2	6	3	20	4	17	13	18	59
OB (51)	33	40	50	51	38	na	39	12	3	na
OP (49)	26	2	32	1	3	na	30	20	12	na
PN (49)	1	11	8	9	2	na	45	21	22	na
PS (67)	4	1	51	44	45	2	25	20	17	na
RN (80)	11	9	5	25	4	8	7	12	21	14
RS (83)	20	73	1	64	3	65	17	31	14	21
SY (65)	2	1	4	38	3	8	20	11	6	9
VS (76)	1	5	10	3	34	19	12	7	22	na
NN (78)	1	66	43	4	40	na	48	9	21	na
OT (77)	15	2	3	73	67	5	58	8	12	4

Nota: Ver Anexo I para visualizar as siglas correspondentes às doenças

Como primeira observação deve referir-se que para os 10 hospitais considerados, somente 4 são classificados para todas as doenças (H2, H9, H10 e H13). Na situação contrária encontram-se os hospitais 86 e 15, os quais foram respectivamente classificados somente em 16 e 19 doenças.

Conforme foi referido os hospitais 9, 13, 10 e 61 são os que recebem mais citações para a melhor efectividade global, para os casos cirúrgicos e médicos e por doença, enquanto que nos hospitais 9, 13, 2 e 61 tal acontece quando se consideram as cinco primeiras posições na efectividade.

No entanto, mesmo para este grupo de hospitais, existe uma grande variabilidade no seu desempenho para cada uma das perspectivas em análise. Utilizando o coeficiente de variação entre as diversas ordenações pode afirmar-se que os hospitais 41, 32 e 48 são os que apresentam uma maior homogeneidade, enquanto que nos hospitais 9, 13 e 2 sucede precisamente o contrário.

A título perfeitamente exemplificativo podem referir-se, por hospital, as seguintes situações como representativas de pior efectividade:

- Hospital 9 – Doenças do sistema nervoso central e dermatológicas;
- Hospital 13 – Casos cirúrgicos e Doenças do sistema musculoesquelético;
- Hospital 61 – Casos obstétricos;
- Hospital 10 – Casos obstétricos e Outras doenças;
- Hospital 2 – Outras doenças;
- Hospital 15 – Doenças da mama e dos órgãos genitais masculinos;
- Hospital 48 – Doenças do sistema nervoso periférico e Imunológicas;
- Hospital 32 – Doenças da mama.

Estas situações reforçam o referido anteriormente, designadamente na necessidade de se melhorar a gestão e a regulação das organizações de saúde, visto que mesmo para os hospitais com melhor efectividade em Portugal são observadas grandes diferenças no desempenho interno.

Por sua vez, estas e outras situações ainda que menos evidentes, devem ser aprofundadas no futuro, tendo em vista o conhecimento das respectivas causas e a melhoria do funcionamento dos hospitais em Portugal.

No Quadro resumo LXIII são identificados os desempenhos dos 10 piores hospitais para todas as admissões, nos casos cirúrgicos e médicos e nas doenças, categorizadas pelo Grande Agrupamento de Doenças do Disease Staging.

Enquanto que o hospital 29 tem doentes tratados em todas as doenças e nos hospitais 33, 40 e 55 tal somente não ocorre em 1 dos 24 Grandes Agrupamento de Doenças, no hospital 88 somente são consideradas 19 doenças.

O hospital 88 é o que recebe mais citações para o pior desempenho, sendo seguido pelos hospitais 26 e 62, ambos com 3 citações.

Quando se consideram as cinco últimas posições na efectividade, com pior desempenho aparece o hospital 59, com 12 citações em 21 possíveis. O hospital 12 recebe 12 citações em 22 doenças e os hospitais 88, 26 e 62 recebem 8, 7 e 7 citações, respectivamente.

No entanto, para os hospitais com pior desempenho global a variabilidade entre casos cirúrgicos e médicos e entre doenças não é tão grande como a verificada para os hospitais com melhor desempenho.

De facto, o hospital 88 é o que apresenta um maior coeficiente de variação (0.50), sendo seguido pelo hospital 33 (com um coeficiente de variação de 0.49). Na situação contrária encontram-se os hospitais 30 e 59, na qual esta estatística é de 0.29 e de 0.31, respectivamente.

Quadro LXIII
Efectividade Global, 10 Piores Hospitais
Desempenho para os Casos Cirúrgicos e Médicos e por Doença
Ordenação

	H12	H59	H62	H88	H30	H55	H43	H40	H33	H29
Total (88)	88	87	86	85	84	83	82	81	80	79
Cirúrgicos (86)	85	73	63	14	75	70	80	84	77	62
Médicos (88)	88	87	85	86	83	84	82	78	79	81
BR (56)	27	52	54	56	40	12	36	9	47	51
CN (84)	84	72	71	39	81	63	82	76	38	79
CV (77)	77	76	50	na	71	70	18	35	74	64
DE (82)	80	55	61	75	76	13	67	78	22	64
EN (63)	38	55	48	63	53	46	12	54	25	34
GI (83)	82	81	80	78	77	83	72	73	76	48
GY (69)	46	49	62	69	52	40	5	55	26	67
HB (80)	76	78	72	69	62	79	73	80	31	41
HE (57)	47	56	33	57	53	50	15	44	51	48
IM (24)	17	na	19	na	na	na	Na	7	11	15
LY (21)	na	na	na	21	na	19	Na	na	na	16
MG (71)	64	67	51	70	38	31	29	8	22	59
MS (82)	81	9	46	71	77	72	68	75	40	57
ND (82)	78	81	82	63	80	76	74	69	79	75
OB (51)	na	24	20	na	14	6	18	5	17	46
OP (49)	10	na	na	42	45	31	40	47	36	13
PN (49)	37	33	32	na	na	34	40	42	14	44
PS (67)	21	66	67	na	58	61	65	35	30	63
RN (80)	71	78	77	79	70	76	72	73	80	66
RS (83)	60	79	83	26	58	82	68	70	74	59
SY (65)	62	53	39	64	47	60	30	63	61	59
VS (76)	75	73	64	4	65	50	70	68	71	67
NN (78)	32	74	64	14	47	44	41	63	59	27
OT (77)	70	71	65	68	74	63	48	75	69	72

Nota: Ver Anexo I para visualizar as siglas correspondentes às doenças

Mesmo assim, existem para este grupo de hospitais situações com bom desempenho, designadamente:

- As Doenças oftalmológicas para os hospitais 12 e 29;
- As Doenças do sistema musculoesquelético para o hospital 59;
- Os Casos obstétricos para os hospitais 30, 40, 43, 55 e 62;
- Os Casos cirúrgicos e as Doenças vasculares e Doenças não especificadas para o hospital 88;
- As Doenças ginecológicas, as Doenças dos olhos, do nariz e da garganta e as Cardiovasculares para o hospital 43;
- As Doenças da mama, Imunológicas e dos Órgãos genitais masculinos para o hospital 40;
- As Doenças dermatológicas para o hospital 55.

Estes aspectos tornam a evidenciar a existência de diferenças na efectividade intra-hospitalar, pelo que o reforço da gestão e dos mecanismos de regulação das organizações de saúde deve passar a constituir uma das prioridades da política de saúde em Portugal.

Para melhor se visualizar a discrepância no desempenho interno dos hospitais em termos de efectividade global e para os casos cirúrgicos e médicos e por doença, foram identificados os hospitais que aparecem mais vezes com citações nestas perspectivas (ver Quadro LXIV).

Quadro LXIV
Hospitais com Maiores Diferenças na Efectividade Global
para Casos Cirúrgicos e Médicos e por Doença

Hospital	Desemp. Global Ordenação	Desemp. Cirúrgicos Ordenação	Desemp. Médicos Ordenação	Doenças Melhor	Doenças 5 Melhores	Doenças Pior	Doenças 5 Piores
1	78	86	61	0	6	1	5
4	21	16	26	2	3	1	3
10	4	9	3	2	8	1	3
13	2	82	2	3	12	0	2
25	60	10	66	4	8	4	4
26	61	53	60	0	2	3	7
40	81	84	78	0	1	1	5
48	7	2	12	0	4	0	2
61	3	3	4	2	9	0	1
86	10	5	13	0	5	0	1
87	73	4	80	1	2	1	6
88	85	14	86	0	6	1	8

Para os 12 hospitais onde foram encontradas maiores diferenças na efectividade por doença podem fazer-se as seguintes observações:

- Existem 5 hospitais entre os 10 com melhor desempenho global, 2 entre os 10 com pior desempenho global, pertencendo os 5 restantes a níveis intermédios de efectividade;
- Destes 12 hospitais, 5 são do Tipo III, 4 do Tipo I e 3 do Tipo V, estando 5 localizados na Região de Saúde IV, 4 na Região III, 2 Na Região III e 1 na Região II;
- Em relação à dimensão da produção verifica-se que 8 hospitais pertencem ao grupo com maior volume e 3 ao 2º grupo em termos de volume. O restante está incluído no 3º grupo de maior volume de produção;
- Os hospitais 1 e 4 são os que simultaneamente apresentam maiores frequências entre os 5 melhores e piores hospitais, na efectividade por doença. Nenhum destes hospitais está incluído entre os melhores ou piores na efectividade global;
- O hospital 13, embora ocupe a 2ª posição na efectividade global e apareça com 12 citações entre os 5 melhores por doença e apenas receba 2 para o pior desempenho, está classificado entre os 10 piores na efectividade para os casos cirúrgicos.

Estes elementos permitem afirmar que praticamente não existe nenhum padrão para a existência de discrepâncias no desempenho interno dos hospitais, no que respeita à efectividade.

Contudo, algumas reflexões suplementares devem ser feitas, nomeadamente:

- Todos os hospitais do Tipo V apresentam simultaneamente níveis de efectividade altos e baixos;
- A dimensão da produção parece ter alguma relevância para a existência de maiores clivagens na efectividade por doença. Este aspecto, só por si, justifica uma análise mais aprofundada, visto que interessa apurar se constitui um fenómeno estrutural, ou se, antes pelo contrário, estará somente a ser um ponto de chegada, na qual a dimensão do hospital poderá ser um “proxy” de maiores problemas na gestão destas organizações de saúde.

6.2.2.2. Eficiência

No capítulo Resultados foram identificados alguns aspectos dentro dos quais se deve destacar:

- Existem indicações completamente distintas quando se analisa o desempenho (eficiência) dos hospitais em função dos seus resultados brutos e ajustados pelo risco. Estas conclusões são válidas para qualquer das perspectivas de análise, ou seja, todas as admissões, casos cirúrgicos e médicos e por doença;
- Quando se pretende ajustar pelo risco existem, mais uma vez, indicações distintas entre os métodos directo e indirecto para a sua concretização, sendo estas conclusões igualmente válidas para todas as perspectivas de análise;
- Observou-se ainda um comportamento bastante heterogéneo dos hospitais quando a avaliação do desempenho é feita de uma forma mais desagregada. Ou seja, os hospitais em termos genéricos apresentam comportamentos distintos entre os casos cirúrgicos e médicos e, de uma forma mais nítida, por doença.

Estas conclusões são praticamente idênticas às encontradas quando se analisa a efectividade dos hospitais, reforçando assim os argumentos apresentados pelos hospitais, os quais parecem ter cabimento, visto que os seus resultados brutos não devem ser comparados, essencialmente devido às características dos seus doentes, onde avulta a gravidade.

Na realidade, alguns hospitais apresentam demora média mais elevada, simplesmente pelo facto da gravidade média dos doentes tratados ser mais elevada. Bons exemplos ilustrativos desta situação são os hospitais 6, 7, 11, 14, 15, 73, 75, 81, 83 e 87.

Na situação contrária, isto é organizações de saúde com bons resultados brutos, encontram-se os hospitais 18, 19, 20, 21, 28, 41, 47, 58, 69 e 85. Assim, estes hospitais aparecem incorrectamente identificados com bom desempenho, sendo a demora média baixa devida à menor gravidade dos doentes tratados.

Outra forma de ilustrar este comportamento deriva da comparação entre os 10 hospitais com demora média mais elevada, com os seus resultados ajustados pelo risco, visto que nenhum deles ocupa as 10 últimas posições na eficiência, embora os hospitais 14 e 75 ocupem respectivamente a 77^a e 74^a posições nesta perspectiva de análise.

Quando se analisam os 10 hospitais com demora média mais baixa, somente os hospitais 32 e 41 aparecem igualmente entre os 10 hospitais com melhor desempenho.

Mais uma vez é observado que os métodos directo e indirecto para se ajustarem os resultados pelo risco conduzem igualmente a indicações diferentes sobre o desempenho dos hospitais.

Os hospitais 68, 13, 17, 78 e 81 são os que aparecem mais beneficiados pelo ajustamento directo (“z score”), enquanto que os hospitais 9, 10, 69, 77 e 65 são os mais prejudicados por este tipo de ajustamento.

Mesmo assim, os dois métodos concordam em 6 hospitais entre os 10 com pior desempenho, mas somente em 3 para os 10 hospitais com melhor desempenho.

Por outro lado, a utilização do método indirecto, conduz a que para os melhores hospitais 5 são do Tipo III, 2 dos Tipos I e IV e 1 do Tipo II. Para o pior desempenho 4 são dos Tipos III e IV e 1 dos Tipos I e II.

Com o método directo encontram-se 7 hospitais do Tipo III e 3 do Tipo I, tanto para o melhor como para o pior desempenho.

Estes resultados, para além dos argumentos teóricos justificativos para a utilização do método directo para se analisar a actividade dos hospitais (Shwartz, Ash e Iezzoni, 1997), evidenciam a existência de hospitais do mesmo Tipo com bons e maus níveis de eficiência, o que para além de tornar inútil a comparação dos hospitais baseada em critérios de oferta, reforça ainda a necessidade de se utilizar o ajustamento pelo risco como metodologia para se avaliar o desempenho dos hospitais.

No que se refere à diferença na eficiência entre os casos médicos e cirúrgicos podem fazer-se as seguintes observações:

- Os hospitais 64, 85, 23, 20 e 54 são os que apresentam maiores discrepâncias no desempenho entre os casos cirúrgicos e os médicos, com melhor desempenho nos primeiros;
- Com melhor desempenho nos casos médicos do que nos cirúrgicos, as maiores discrepâncias são encontradas nos hospitais 31, 35, 87, 24 e 49;
- Por sua vez, os hospitais 5, 21, 40, 41, 60 e 79 são os que apresentam um desempenho idêntico entre os casos médicos e os cirúrgicos.

Para os 5 hospitais que comparativamente apresentam melhor desempenho nos casos cirúrgicos do que nos médicos, não existe um padrão na eficiência global, visto que o hospital 85 é o que aparece mais bem classificado (18^a posição) e o hospital 54 o pior colocado (70^a posição).

Para os 5 hospitais que comparativamente apresentam melhor desempenho nos casos médicos, a situação é a mesma, visto que o hospital 49 ocupa a 25^a

posição para todas as admissões, enquanto que o hospital 31 ocupa a 59ª posição.

Assim, para os 6 hospitais que apresentam desempenhos semelhantes entre os casos médicos e cirúrgicos não se vislumbra nenhum padrão. Enquanto que os hospitais 41, 21 e 5 têm bom desempenho para todas as admissões, os hospitais 60 e 40 têm mau desempenho para todas as admissões.

Para além destes aspectos existem outras questões que merecem uma discussão suplementar, designadamente os relacionados com algumas questões teóricas e outros com potenciais contradições encontradas neste estudo.

Para as primeiras (questões conceptuais) à semelhança do realizado para a efectividade, serão analisados alguns tópicos referenciados na literatura internacional, como relevantes para diferentes desempenhos dos hospitais. Destes destacam-se o tipo de hospital, a localização do hospital e a dimensão do hospital (Wagstaff, 1989; Paiva, 1993; Butler, 1995; Given, 1996; Wholey et al, 1996; Dranove, 1998; Barros e Sena, 1999; Cabral e Barriga, 1999 e Carreira, 1999).

Para as questões relacionadas com os resultados deste estudo, avultam as relativas a eventuais desempenhos distintos entre os hospitais, tanto para casos cirúrgicos e médicos, como para diferenças entre as doenças.

Em relação às questões conceptuais serão feitas análises de regressão em que as variáveis dependentes são os resultados brutos (demora média) e os resultados ajustados pelo risco, pelo método directo (“z score”).

As variáveis explicativas são as seguintes:

- Dimensão – medida pelo volume total da produção por hospital (doentes saídos). Para se pesquisarem eventuais relações não lineares, será igualmente utilizada produção elevada ao quadrado;
- Tipo de hospital – será utilizada uma variável “dummy”, identificando os seguintes tipos de hospital: Central e Universitário; Especializado; Distrital; Distrital de Nível 1 e Institutos Portugueses de Oncologia (IPOs);
- Localização do hospital – será utilizada uma variável “dummy”, identificando os hospitais em função da respectiva Região de Saúde: Norte, Centro, Lisboa e Vale do Tejo, Alentejo e Algarve. A este respeito, embora a distinção entre a localização urbana e rural também seja relevante, optou-se por esta operacionalização, para existir uma maior correspondência com a lógica de gestão em vigor em Portugal.

Para estas análises de regressão serão analisados o total das admissões e os casos cirúrgicos e médicos.

Os principais resultados são os seguintes:

Total das Admissões

Resultados brutos:

R2 (ajustado) = 0.225; Estatística F = 3.520 (significativa a 0.05).

Figura 65
Resultados do Modelo: Demora Média; Total das Admissões

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Sim
Produção (termo ao quadrado)	Sim
Centrais e Universitários	Sim
Especializados	Sim
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Não

A demora média dos hospitais está a ser explicada pela produção, sendo mais elevada nas organizações com mais e menos produção (distribuição em forma de U) e pelo Tipo de hospital, embora só seja significativo o coeficiente de regressão para os hospitais Centrais e Universitários e Especializados. Deve ainda referir-se que não existe multicolinearidade no modelo, visto que o “Eigenvalue” é de 0.18, com um Índice de Condição de 13.794 e que os resíduos são normais com a estatística do Teste de Kolmogorov-Smirnov a ter o valor de 1.033 (probabilidade de significância > 0.05).

Estes resultados parecem estar de acordo com o padrão esperado para o Tipo de hospital, visto que a demora média é maior nos hospitais centrais e universitários e menor nos especializados.

Contudo, para a produção a existência de uma relação quadrática com a demora média, a qual, por sua vez, traduz uma distribuição em forma de U, permite afirmar a existência de uma dimensão óptima, a partir da qual a demora média passa a aumentar, tanto para valores superiores, como inferiores.

Analisando agora os resultados ajustados pelo risco os valores são os listados de seguida.

Os resultados ajustados pelo risco não estão a ser explicados pelo modelo, sendo esta conclusão bem expressa pelo valor do R2 ajustado e pela estatística F.

As mesmas indicações são encontradas, variável a variável, dado que nenhum dos coeficientes de regressão é significativo.

R2 (ajustado) = -0.024; Estatística F = 0.797 (não significativa)

Figura 66
Resultados do Modelo: Eficiência; Total das Admissões

Variáveis explicativas	Significância (< 0.05)
Constante	Não
Produção (termo simples)	Não
Produção (termo ao quadrado)	Não
Centrais e Universitários	Não
Especializados	Não
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Não

Somente a título indicativo, deve referir-se que não existe multicolinearidade no modelo, sendo os valores do “Eigenvalue” e do Índice de Condição, iguais aos do modelo anterior. Os resíduos são normais, assumindo a estatística do Teste de Kolgomorov-Smirnov um valor de 1.067, sendo a probabilidade de significância > 0.05.

Estes resultados do modelo de regressão permitem concluir que a eficiência dos hospitais não está a ser explicada pela produção, nem pelos critérios de oferta utilizados, o Tipo de hospital e a sua localização.

Dado que estas indicações são diferentes das encontradas para a efectividade, no capítulo seguinte será feita uma análise, tendo em vista a discussão se existem contradições ou conflitos entre a efectividade e a eficiência dos hospitais.

Casos Cirúrgicos

Resultados brutos:

R2 (ajustado) = 0.385; Estatística F = 6.322 (significativa a 0.05).

Figura 67
Resultados do Modelo: Demora Média; Casos Cirúrgicos

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Não
Produção (termo ao quadrado)	Não
Centrais e Universitários	Sim
Especializados	Não
Distritais de Nível 1	Sim
IPOs	Sim
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Sim

A demora média está somente a ser explicada pelo Tipo de hospital (todos os coeficientes de regressão são significativos, com exceção dos hospitais especializados) e pela sua localização (sendo somente significativa a Região Norte). Mais uma vez não existe multicolinearidade no modelo e os resíduos são normais.

Para os resultados ajustados pelo risco os valores são os seguintes:

R^2 (ajustado) = 0.179; Estatística F = 2.855 (significativa a 0.05).

Figura 68
Resultados do Modelo: Eficiência; Casos Cirúrgicos

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Sim
Produção (termo ao quadrado)	Sim
Centrais e Universitários	Sim
Especializados	Não
Distritais de Nível 1	Sim
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Não

Não existe multicolinearidade no modelo e os resíduos são normais, mas a eficiência para os casos cirúrgicos é somente explicada pelo nível produtivo e pelo Tipo de hospitais (somente os hospitais centrais e universitários e de nível 1).

Para os casos cirúrgicos os resultados da análise de regressão são ainda mais surpreendentes.

Em primeiro lugar, porque a dimensão de produção não tem efeitos sobre a demora média, mas o mesmo não sucede com a eficiência, com uma distribuição em U.

Em segundo lugar, ao contrário do que sucedeu para a efectividade onde se encontrou igualmente um efeito dimensão produtiva, mas desta vez numa forma de U invertido, embora a maior efectividade seja encontrada nos hospitais com maior produção, na eficiência os hospitais com piores resultados são os com maior produção.

Em terceiro lugar, existe um comportamento diferente do verificado para o total das admissões, onde as variáveis consideradas não explicam, a eficiência.

Os hospitais 41, 26, 25, 1, 48, 5, 59, 51, 13 e 32 são os que apresentam melhor eficiência do que a esperada em função da sua produção e tipo de hospital. Estes hospitais são 6 do Tipo III e 4 do Tipo I, estando 5 localizados na Região de Saúde IV, 3 na Região III e 2 na Região V.

Para os 10 hospitais com valores melhores que o esperado verifica-se que, com excepção do hospital 51, já ocupavam as 20 primeiras posições em termos de eficiência.

Os hospitais 12, 61, 60, 9, 4, 29, 27, 44, 55 e 31 são os que apresentam pior efectividade face à esperada em função da sua produção e tipo de hospital. Encontram-se 7 hospitais do Tipo III e 3 do Tipo I, estando 4 localizados na Região de Saúde IV e 3 nas Regiões III e V.

Mais uma vez são encontradas indicações eventualmente diferentes entre a efectividade e a eficiência, aspectos que serão discutidos no capítulo seguinte.

Casos Médicos

O modelo não apresenta multicolinearidade e os resíduos são normais, sendo os resultados brutos exclusivamente explicados pela dimensão da produção (a demora média diminui à medida que aumenta a produção) e pelo tipo de hospital (somente os centrais e universitários e os especializados apresentam coeficientes de regressão significativos).

Resultados brutos:

R^2 (ajustado) = 0.268; Estatística F = 4.185 (significativa a 0.05).

Figura 69
Resultados do Modelo: Demora Média; Casos Médicos

Variáveis explicativas	Significância (< 0.05)
Constante	Sim
Produção (termo simples)	Sim
Produção (termo ao quadrado)	Não
Centrais e Universitários	Sim
Especializados	Sim
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Não

Para os resultados ajustados pelo risco os valores são os seguintes:

R2 (ajustado) = -0.011; Estatística F = 0.908 (não significativa).

Figura 70
Resultados do Modelo: Eficiência; Casos Médicos

Variáveis explicativas	Significância (< 0.05)
Constante	Não
Produção (termo simples)	Não
Produção (termo ao quadrado)	Não
Centrais e Universitários	Não
Especializados	Não
Distritais de Nível 1	Não
IPOs	Não
Alentejo	Não
Algarve	Não
Centro	Não
Norte	Não

Mais uma vez, para os casos médicos, os resultados ajustados pelo risco diferem dos resultados brutos, visto que a eficiência não é explicada pelo modelo, enquanto que a demora média (resultados brutos) é explicada pela produção, com uma relação inversa e pelo tipo de hospital.

Em síntese pode afirmar-se:

- Existe um efeito dimensão da produção para os resultados ajustados pelo risco somente para os casos cirúrgicos;

- Para estas situações o ajustamento que melhor traduz esta relação assume uma forma quadrática, em U, pelo que a melhor eficiência é encontrada nos hospitais com dimensões médias;
- Para o total das admissões e para os casos médicos, as variáveis incluídas no modelo, não justificam o desempenho dos hospitais.

Esta análise macro, embora não sendo totalmente conclusiva para justificar os diversos níveis de eficiência, permite desde já identificar os seguintes aspectos:

- Comportamento distinto entre o total das admissões e os casos médicos e os casos cirúrgicos;
- Pouca relevância da dimensão e das características da oferta no desempenho dos hospitais;
- Comportamento distinto entre a efectividade e a eficiência dos hospitais.

Enquanto que, os dois primeiros aspectos relevam os problemas da insuficiência na gestão dos hospitais portugueses, como vem sendo referido ao longo deste estudo, o último carece de análises mais aprofundadas, as quais serão realizadas no capítulo seguinte.

Em relação ao diferente desempenho dos hospitais para casos cirúrgicos e médicos, igualmente retratado pelo valor do *K de Cohen* (-0.012), será feita uma análise semelhante à exercida para a discussão sobre a efectividade.

O coeficiente de correlação entre a eficiência dos hospitais para os casos cirúrgicos e médicos é de 0.510 (significativo a 0.01), pelo que se pode concluir que, na generalidade, a eficiência dos hospitais aumenta nos casos cirúrgicos à medida que aumenta a eficiência dos casos médicos e vice-versa.

O coeficiente de correlação de Spearman para a ordenação dos hospitais por nível de eficiência para os casos médicos e cirúrgicos é de 0.571 (significativo a 0.01), o que mais uma vez parece traduzir uma associação directa na classificação dos hospitais para os casos médicos e cirúrgicos.

Neste sentido, embora os hospitais sejam classificados de forma diferente segundo o seu desempenho para os casos cirúrgicos e médicos, tudo parece indicar que este fenómeno não é estrutural, visto que, na generalidade, os hospitais que são mais eficientes nos casos cirúrgicos, também o são nos casos médicos.

No Quadro LXV é apresentada a matriz da ordenação dos hospitais por eficiência para os casos cirúrgicos e médicos por quartis.

Assinalados a azul estão os 37 hospitais que estão classificados no mesmo grupo de desempenho. A vermelho estão identificados 15 hospitais com maiores

discrepâncias no desempenho entre os casos cirúrgicos e os médicos. Para os restantes 34 hospitais considera-se que as diferenças não são muito relevantes.

Quadro LXV
Eficiência
Ordenação do Desempenho dos Hospitais, por Quartis
Casos Cirúrgicos e Casos Médicos

Casos Cirúrgicos	Casos Médicos			
	Até 1º Quartil	Entre o 1º e o 2º Quartis	Entre o 2º e o 3º Quartis	Depois do 3º Quartil
Até 1º Quartil	14	4	3	1
Entre o 1º e o 2º Quartis	4	5	8	4
Entre o 2º e o 3º Quartis	3	9	5	4
Depois do 3º Quartil	1	3	5	13

Esta análise vem confirmar as indicações retiradas da análise de correlação, visto que 37 hospitais (43.02%) apresentam o mesmo desempenho para os casos cirúrgicos e médicos, enquanto que somente 17.44.% (15 hospitais) apresentam desempenhos completamente distintos para estes dois grandes conjuntos de doentes.

Para se analisar uma eventual existência de um padrão para o diferente comportamento dos hospitais para os casos médicos e cirúrgicos, serão utilizados os três grupos de hospitais: os hospitais com menores discrepâncias entre os casos cirúrgicos e médicos, os hospitais com discrepâncias médias nestes mesmos casos e os hospitais com maiores discrepâncias.

Estes grupos serão testados em função dos seguintes aspectos:

- Desempenho dos hospitais;
- Tipo do hospital;
- Localização do hospital;
- Produção do hospital;
- “Mix” de casos tratados.

Em relação ao desempenho dos hospitais (ver elementos do Quadro LXV) observa-se que para os 37 hospitais que apresentam o mesmo comportamento nos casos cirúrgicos e médicos, uma distribuição bastante diferente por níveis de desempenho, visto que 14 se encontram entre os melhores hospitais, 13 entre os piores e 5 para cada um dos níveis intermédios de desempenho (eficiência).

No entanto, para os 15 hospitais com maiores discrepâncias no desempenho entre os casos médicos e cirúrgicos a situação é diferente, visto que existe uma distribuição praticamente idêntica por níveis de eficiência.

Neste momento e atendendo somente ao desempenho global dos hospitais, pode afirmar-se que parece não existir um padrão justificativo para o diferente desempenho dos hospitais para os casos cirúrgicos e médicos.

No Quadro LXVI são apresentados os hospitais segundo a heterogeneidade do desempenho entre os casos cirúrgicos e médicos por tipo de hospital. Entre parênteses encontram-se os valores esperados se a distribuição fosse equitativa por todas as células.

Quadro LXVI
Eficiência
Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos
por Tipo de Hospital

	Tipo I	Tipo II	Tipo III	Tipo IV	Tipo V
Pequenas Variações	9 (6)	1 (3)	18 (17)	8 (9)	1 (1)
Variações Médias	5 (6)	3 (2)	15 (16)	11 (9)	0 (1)
Grandes Variações	1 (3)	2 (1)	7 (7)	3 (4)	2 (1)

Existem menos 3 hospitais do Tipo I que os esperados para o grupo das pequenas variações e menos 2 que os esperados, ainda para o mesmo tipo de hospitais no grupo das grandes variações. A outra diferença que se deve referir é encontrada nos hospitais do Tipo IV (mais 2 hospitais que o esperado) para o grupo das variações médias.

Esta situação permite afirmar que não parecem existir grandes diferenças entre os valores observados e esperados, pelo que à semelhança do que foi referido para a análise da efectividade, não parece existir nenhum padrão em função do tipo de hospital, para justificar o diferente comportamento destas organizações para os casos cirúrgicos e médicos.

No Quadro LXVII são apresentados os hospitais segundo a heterogeneidade do desempenho entre os casos cirúrgicos e médicos por Região de Saúde. Entre parênteses encontram-se os valores esperados se a distribuição fosse equitativa por todas as células.

A análise regional vem evidenciar a existência de algumas diferenças significativas, designadamente:

- Para a Região V existem mais 4 hospitais do que o esperado no grupo das grandes variações e igualmente menos 4 hospitais que o esperado no grupo das variações médias;
- Na Região III existem menos 3 hospitais que o esperado no grupo das grandes variações e mais 3 que o esperado no grupo das variações médias;
- Na Região I existem mais 2 hospitais que o esperado no grupo das variações médias.

Quadro LXVII
Eficiência
Diferenças no Desempenho dos Hospitais entre Casos Cirúrgicos e Médicos
por Região de Saúde

	Região I	Região II	Região III	Região IV	Região V
Pequenas Variações	1 (2)	2 (1)	11 (11)	12 (12)	11 (11)
Variações Médias	4 (2)	1 (1)	13 (10)	10 (11)	6 (10)
Grandes Variações	0 (1)	0 (1)	2 (5)	5 (5)	8 (4)

Estes aspectos, pese embora o facto de nas Regiões de Saúde II e IV não se terem encontrado comportamento distinto dos hospitais na eficiência para os casos cirúrgicos e médicos, permitem identificar a Região V como aquela em que os hospitais têm desempenhos mais distintos entre os casos cirúrgicos e os médicos.

Assim, pode concluir-se que a existência de um padrão regional para a eficiência nos casos cirúrgicos e médicos não é totalmente patente, embora a situação encontrada na Região V exija a realização futura de estudos mais detalhados para melhor caracterizar este fenómeno.

No Quadro LXVIII são apresentadas a produção média e respectivo coeficiente de variação, e a percentagem de casos cirúrgicos, para todos os hospitais, bem como para os grupos de variação entre casos cirúrgicos e médicos anteriormente definidos.

Quadro LXVIII
Produção Média e Coeficiente de Variação; % de Casos Cirúrgicos
Todos Hospitais e por Grupos de Variação

	Produção	Coeficiente de Variação	% Casos Cirúrgicos
Hospitais	25230	0.88	38.6
Pequenas Variações	31220	0.88	38.1
Variações Médias	21219	0.88	37.2
Grandes Variações	19546	0.35	39.7

Nos hospitais com maior heterogeneidade na eficiência entre os casos cirúrgicos e médicos observa-se uma dimensão produtiva inferior à da média dos hospitais, bem como à dos restantes grupos de variação. A heterogeneidade produtiva neste grupo de hospitais, medida pelo coeficiente de variação é a mais reduzida de todas as perspectivas de análise.

No grupo de variações médias na efectividade entre os casos cirúrgicos e médicos observa-se que a produção média é a 2^a mais baixa, embora a heterogeneidade dos hospitais dentro do grupo seja igual à dos restantes grupos, com excepção do grupo com maior variação.

No grupo com maior homogeneidade na efectividade entre os casos cirúrgicos e médicos observa-se a maior produção média, com uma variação entre os hospitais do grupo semelhante à dos restantes grupos.

A análise de variância simples (“one-way anova”) permite afirmar que as diferenças na produção média dos grupos de hospitais não são significativas, situação que se mantém quando se analisam os grupos de variação dois a dois, pelo que se pode concluir que a produção não tem influência no desempenho distinto dos hospitais entre os casos médicos e cirúrgicos.

Recorde-se a este propósito que na análise da efectividade a produção tinha alguma influência nas diferenças encontradas entre hospitais para casos cirúrgicos e médicos, pelo que mais uma vez parecem ser encontrados comportamentos distintos dos hospitais para a efectividade e para a eficiência.

Finalmente deve referir-se que, embora a percentagem de casos cirúrgicos tratados seja similar, para os três grupos de hospitais em análise, existem mais casos cirúrgicos no grupo de hospitais com maiores diferenças na eficiência entre os casos cirúrgicos e médicos.

No capítulo Resultados foram evidenciados alguns aspectos, quando se analisa a eficiência dos hospitais por doença, nomeadamente a existência de distintos hospitais com melhor eficiência por doença, tanto para a 1^a posição, como para as cinco primeiras posições ou ainda para os hospitais com pior efectividade (igualmente para a última posição, ou para as cinco últimas posições) e uma grande heterogeneidade na efectividade dentro de cada hospital em função das diversas doenças.

Estas questões conduzem à necessidade de reforço da gestão e dos mecanismos de regulação dos hospitais e suscitam novos problemas para o financiamento das organizações de saúde e para o exercício da liberdade de escolha por parte dos consumidores.

Em seguida, será analisada a eficiência dos hospitais por doença, tendo como cenário de fundo a existência de grandes diferenças intra-hospitalares.

Para melhor se visualizar este fenômeno será apresentado para os 10 hospitais com melhor eficiência global, para os 10 com pior eficiência global, bem como para as organizações com maior heterogeneidade na eficiência, o resumo do seu desempenho para os casos cirúrgicos e médicos e por doença.

No Quadro LXIX são identificados os desempenhos dos 10 melhores hospitais para todas as admissões, nos casos cirúrgicos e médicos e nas doenças, categorizadas pelo Grande Agrupamento de Doenças do Disease Staging.

Quadro LXIX
Eficiência Global, 10 Melhores Hospitais
Desempenho para os Casos Cirúrgicos e Médicos e por Doença
Ordenação

	H41	H25	H13	H26	H48	H32	H16	H6	H59	H58
Total (88)	1	2	3	4	5	6	7	8	9	10
Cirúrgicos (86)	1	2	13	3	4	15	21	16	7	8
Médicos (88)	2	5	1	13	19	4	3	6	20	16
BR (56)	4	2	6	11	9	12	1	21	13	25
CN (84)	6	5	1	63	14	7	4	34	16	3
CV (77)	11	7	6	61	20	9	12	10	23	64
DE (82)	4	11	74	1	7	16	6	23	37	32
EN (63)	2	11	42	32	3	15	21	na	39	5
GI (83)	3	4	22	1	10	7	18	6	36	35
GY (69)	9	32	6	48	3	24	7	27	19	20
HB (80)	4	6	10	3	8	7	36	1	64	16
HE (57)	15	34	49	14	42	na	6	4	12	47
IM (24)	na	20	11	21	10	na	1	6	na	na
LY (21)	na	12	20	15	na	na	Na	8	na	na
MG (71)	4	33	23	3	5	18	21	61	9	19
MS (82)	2	1	5	25	8	20	13	54	3	7
ND (82)	8	1	2	3	9	5	7	79	37	24
OB (51)	4	30	38	1	14	39	23	na	15	9
OP (49)	4	36	7	31	22	24	Na	na	na	na
PN (49)	16	38	41	1	20	15	Na	2	23	21
PS (67)	16	28	6	2	18	12	47	14	19	51
RN (80)	12	8	13	53	16	6	15	18	24	26
RS (83)	5	3	7	40	26	4	22	1	24	49
SY (65)	21	6	2	50	17	4	8	9	13	41
VS (76)	6	5	4	7	12	10	37	49	54	26
NN (78)	13	35	9	24	38	17	5	18	31	4
OT (77)	6	39	43	64	7	25	9	29	66	54

Nota: Ver Anexo I para visualizar as siglas correspondentes às doenças

Na primeira coluna, entre parênteses está identificado o número de hospitais que foram classificados em cada perspectiva de análise. A este propósito, recorde-se que não são considerados os hospitais que trataram menos de 100 doentes.

Como primeira observação deve referir-se que para os 10 hospitais considerados, somente 3 são classificados para todas as doenças (H13, H25 e H126). Na situação contrária encontram-se os hospitais 6, 16, 32, 58 e 59, os quais foram classificados somente em 21 doenças.

Conforme foi referido o hospital 26 é o que apresenta mais citações na eficiência por doença, enquanto que nos hospitais 41, 26 e 25 tal acontece quando se consideram as cinco primeiras posições na eficiência.

No entanto, mesmo para este grupo de hospitais, existe uma grande variabilidade no seu desempenho para cada uma das perspectivas em análise. Utilizando o coeficiente de variação entre as diversas ordenações pode afirmar-se que os hospitais 32, 59 e 58 são os que apresentam uma maior homogeneidade, enquanto que nos hospitais 13, 26 e 6 sucede precisamente o contrário.

A título perfeitamente exemplificativo podem referir-se, por hospital, as seguintes situações como representativas de pior efectividade:

- Hospital 13 – Doenças dermatológicas, Hematológicas e do Sistema nervoso periférico;
- Hospital 26 – Doenças do sistema nervoso central, Cardiovasculares e Outras;
- Hospital 48 – Doenças hematológicas;
- Hospital 32 – Casos obstétricos;
- Hospital 6 – Doenças dos órgãos genitais masculinos e Endócrinas;
- Hospital 59 – Doenças hepatobiliares e Outras;
- Hospital 68 – Doenças cardiovasculares e Hematológicas.

Estas situações reforçam o referido anteriormente, designadamente na necessidade de se melhorar a gestão e a regulação das organizações de saúde, visto que mesmo para os hospitais com melhor eficiência em Portugal são observadas grandes diferenças no desempenho interno.

Por sua vez, estas e outras situações ainda que menos evidentes, devem ser aprofundadas no futuro, tendo em vista o conhecimento das respectivas causas e a melhoria do funcionamento dos hospitais em Portugal.

No Quadro LXX são identificados os desempenhos dos 10 piores hospitais para todas as admissões, nos casos cirúrgicos e médicos e nas doenças, categorizadas pelo Grande Agrupamento de Doenças do Disease Staging.

Quadro LXX
Eficiência Global, 10 Piores Hospitais
Desempenho para os Casos Cirúrgicos e Médicos e por Doença
Ordenação

	H9	H36	H12	H60	H38	H29	H44	H27	H4	H43
Total (88)	88	87	86	85	84	83	82	81	80	79
Cirúrgicos (86)	86	83	85	81	75	65	49	77	82	73
Médicos (88)	76	88	49	84	85	86	87	83	81	82
BR (56)	49	35	50	38	43	45	19	23	48	39
CN (84)	80	78	75	23	74	81	82	21	83	79
CV (77)	3	72	18	68	71	73	40	74	33	49
DE (82)	2	72	79	80	60	75	59	63	69	73
EN (63)	7	37	61	52	43	18	59	53	41	57
GI (83)	81	78	77	80	45	79	76	73	82	65
GY (69)	66	60	55	61	57	11	53	64	56	69
HB (80)	78	74	76	72	31	73	70	65	69	60
HE (57)	3	21	43	39	54	55	57	51	1	40
IM (24)	15	na	16	na	na	na	Na	na	19	na
LY (21)	21	na	na	11	na	17	19	na	10	na
MG (71)	69	68	39	60	43	62	52	70	1	44
MS (82)	82	76	81	73	78	29	72	15	56	18
ND (82)	73	80	78	59	69	66	50	77	76	72
OB (51)	51	42	na	44	41	31	8	47	na	45
OP (49)	3	18	46	16	47	42	9	37	48	43
PN (49)	47	na	32	27	na	33	34	17	45	46
PS (67)	65	38	34	63	58	56	55	59	39	9
RN (80)	80	75	19	68	72	78	41	76	3	32
RS (83)	83	78	50	61	75	76	82	33	44	80
SY (65)	3	58	61	46	52	60	59	53	55	36
VS (76)	3	72	70	67	69	73	21	56	68	45
NN (78)	1	60	28	69	22	25	58	72	65	36
OT (77)	70	67	36	74	69	63	73	71	15	51

Nota: Ver Anexo I para visualizar as siglas correspondentes às doenças

Enquanto que o hospital 9 tem doentes tratados em todas as doenças, nos hospitais 4, 29, 44 e 60 tal somente não ocorre em 1 dos 24 Grandes Agrupamento de Doenças e nos hospitais 36 e 38 somente são consideradas 21 doenças.

O hospital 9 é o que apresenta mais citações para o pior desempenho. Quando se consideram as cinco últimas posições na eficiência, com pior desempenho aparece o hospital 9, com 12 citações em 24 possíveis. O hospital 29 recebe 8 citações em 23 doenças, o hospital 29 com 7 citações em 22 doenças e os hospitais 27, 44 e 12 recebem 7, 6 e 6 citações, respectivamente.

No entanto, para os hospitais com pior desempenho global a variabilidade entre casos cirúrgicos e médicos e entre doenças não é tão grande como a verificada para os hospitais com melhor desempenho.

De facto, o hospital 9 é o que apresenta um maior coeficiente de variação (0.72), sendo seguido pelo hospital 4 (com um coeficiente de variação de 0.56). Na situação contrária encontram-se os hospitais 38 e 36, na qual esta estatística toma os valores de 0.29 e de 0.32, respectivamente.

Mesmo assim, existem para este grupo de hospitais situações com bom desempenho, designadamente:

- No hospital 9 as Doenças cardiovasculares, Dermatológicas, Hematológicas, Oftalmológicas, Multisistémicas, Vasculares e Não especificadas;
- No hospital 4 as Doenças hematológicas, dos Órgãos genitais masculinos e Renais.

Estes aspectos tornam a evidenciar a existência de diferenças na eficiência intra-hospitalar, pelo que o reforço da gestão e dos mecanismos de regulação das organizações de saúde deve passar a constituir uma das prioridades da política de saúde em Portugal.

Para melhor se visualizar a discrepância no desempenho interno dos hospitais em termos de eficiência global e para os casos cirúrgicos e por doença, foram identificados os hospitais que aparecem mais vezes com citações nestas duas perspectivas (ver Quadro LXXI).

Quadro LXXI
Hospitais com Maiores Diferenças na Eficiência Global
para Casos Cirúrgicos e Médicos e por Doença

Hospital	Desemp. Global Ordenação	Desemp. Cirúrgicos Ordenação	Desemp. Médicos Ordenação	Doenças Melhor	Doenças 5 Melhores	Doenças Pior	Doenças 5 Piores
1	11	5	43	2	9	1	5
2	20	33	12	2	5	0	2
4	80	82	81	2	5	0	2
7	39	57	32	0	2	0	2
9	88	86	76	1	7	7	12
10	75	80	71	0	2	2	8
15	26	50	21	1	2	1	2

Para os 7 hospitais onde foram encontradas maiores diferenças na eficiência por doença podem fazer-se as seguintes observações:

- Todos os hospitais são do Tipo I, estando 3 localizados na Região de Saúde IV e 2 nas Regiões III e V;
- Destes 7 hospitais somente 2 estão classificados entre os 10 com pior eficiência global;
- Em relação à dimensão da produção verifica-se que 4 hospitais pertencem ao grupo com maior volume e 3 ao 2º grupo em termos de volume;
- Os hospitais 1 e 9 são os que simultaneamente apresentam maiores frequências entre os 5 melhores e piores hospitais, na eficiência por doença (o hospital 9 é o menos eficiente para todas as admissões);
- O hospital 10, embora não ocupe as últimas posições na eficiência global, aparece 8 vezes entre os 5 piores por doença.

Estes elementos permitem afirmar que praticamente não existe nenhum padrão para a existência de discrepâncias no desempenho interno dos hospitais, no que respeita à eficiência.

Contudo, algumas reflexões suplementares devem ser feitas, nomeadamente:

- Todos os hospitais com grandes discrepâncias internas na eficiência são do Tipo I;
- A dimensão da produção parece ter alguma relevância para a existência de maiores clivagens na eficiência por doença. Este aspecto, só por si, justifica uma análise mais aprofundada, visto que interessa apurar se constitui um fenómeno estrutural, ou se, antes pelo contrário, estará somente a ser um ponto de chegada, na qual a dimensão do hospital poderá ser um “proxy” de maiores problemas na gestão destas organizações de saúde.

6.2.2.3. Desempenho Global

No capítulo Resultados foram apresentadas algumas das questões que devem estar associadas a esta finalidade, designadamente as relacionadas com:

- As dimensões que devem ser consideradas para avaliar o desempenho global dos hospitais;
- As medidas que devem ser utilizadas para avaliar o desempenho dos hospitais;
- A forma de se valorizar as dimensões escolhidas.

O primeiro aspecto pretende responder a um questão simples: o que é que se deve avaliar no desempenho dos hospitais?

Os dois restantes estão relacionados com a metodologia da medição, ou seja, como é que se deve operacionalizar a avaliação do desempenho hospitalar?

Como foi referido, neste estudo embora se tenha presente que a estrutura, o processo e os resultados constituem dimensões importantes para se avaliar o desempenho dos hospitais, somente se utilizam os resultados para concretizar esta finalidade, pelo que apenas são utilizadas as dimensões efectividade e eficiência.

No que respeita à metodologia para se medir o desempenho global dos hospitais foram consideradas diversas alternativas para a sua valorização.

Em primeiro lugar, considerou-se a possibilidade de se utilizarem as ordenações dos hospitais na efectividade e na eficiência ou os valores reais dos hospitais para estas duas dimensões.

Neste estudo provou-se que o desempenho global dos hospitais é diferente em função da medida que é utilizada (por exemplo, para a totalidade das admissões o *K de Cohen* foi de 0.103) e optou-se pela utilização dos valores reais (“z scores” da efectividade e da eficiência) para se avaliar o desempenho global dos hospitais.

Esta situação deve-se essencialmente ao facto de a utilização dos valores reais apresentar uma sensibilidade maior do que a proporcionada pelas ordenações para discriminar o desempenho dos hospitais.

Em seguida, discutiu-se a valorização das dimensões consideradas para avaliação do desempenho global hospitalar, com a construção de três cenários: valorização idêntica, maior peso à efectividade e maior peso à eficiência.

Mais uma vez foi provado que a diferente valorização das dimensões conduz a resultados diferentes na avaliação do desempenho dos hospitais. A este propósito recorde-se que o *K de Cohen* foi de 0.103 entre a valorização idêntica da efectividade e da eficiência e quando se dá maior peso à efectividade e de 0.092 entre a valorização idêntica e quando se dá maior peso à eficiência.

Estes resultados foram analisados, por um lado, face à preferência dos diversos agentes do sector da saúde e, por outro lado, face à existência de evidência científica sobre a melhor metodologia desta valorização.

Neste sentido, pese embora o facto de os proprietários e os gestores darem prioridade à eficiência e de os prestadores e os consumidores à efectividade, optou-se pela valorização idêntica destas duas dimensões para avaliar o desempenho dos hospitais.

Todas estas questões merecem ainda algumas reflexões suplementares:

- A avaliação do desempenho dos hospitais somente em função dos seus resultados é suficiente?
- Qual a melhor metodologia para se valorizarem as dimensões utilizadas no modelo de avaliação do desempenho dos hospitais?

Para o primeiro aspecto, refira-se que mais uma vez pode existir uma contradição entre os diversos agentes do sector da saúde. De facto, enquanto que os proprietários e os gestores podem eventualmente considerar que os resultados são suficientes para se avaliar o desempenho dos hospitais, os prestadores e os consumidores, embora por razões diferentes, podem igualmente valorizar a estrutura e o processo.

Para os consumidores as principais razões justificativas para este comportamento devem ser encontradas na eventual redução na assimetria de informação existente, para que desta forma possam exercer a sua liberdade de escolha.

Para os prestadores, visto que a estrutura pode ser considerada como um “proxy” do desenvolvimento tecnológico dos hospitais e que o processo, para além de poder retratar o estado da arte da prestação de cuidados de saúde, pode ainda dar maior sustentabilidade à evidência clínica, pelo que constituem dimensões que interessa monitorar e valorizar para se apreciar o desempenho dos hospitais.

Neste estudo, embora sem se pretender dar respostas definitivas a este debate, parece ser mais interessante adoptar a cultura organizacional predominante nos EUA, pelo que se considera que os resultados são uma dimensão prioritária para avaliar o desempenho dos hospitais.

No entanto, como foi referido, não deve ser entendido que as dimensões estrutura e processo não são relevantes, mas antes pelo contrário que os resultados constituem uma dimensão necessária, mas não suficiente para se avaliar o desempenho hospitalar, tornando igualmente mais explícito quais os principais problemas existentes em cada hospital.

No que se refere ao segundo aspecto, reitera-se que a utilização dos valores reais e a valorização idêntica da efectividade e da eficiência constituem a melhor metodologia para se medir o desempenho dos hospitais.

Contudo, deve referir-se que atendendo à diversidade dos resultados do desempenho dos hospitais em função dos critérios utilizados, é indispensável que qualquer modelo de classificação dos hospitais explicita a operacionalização da medição do seu desempenho e que, no futuro, se realizem estudos empíricos, tendentes a investigar qual o melhor “mix” para se valorizarem as dimensões que se utilizarão no modelo.

Em função dos critérios utilizados foram identificados os hospitais com melhor desempenho global, com as seguintes particularidades:

- Nos 10 melhores hospitais encontram-se 5 tanto do Tipo I como do Tipo III, sendo 9 comuns com a classificação da efectividade e 5 com a classificação da eficiência;
- Para os 10 piores hospitais encontra-se uma maior concentração pela tipologia, visto que 8 são do Tipo III, sendo os restantes 2 dos Tipos I e V;
- No entanto, na globalidade, os hospitais são classificados de forma diferente, em função da efectividade e da eficiência, dado que a concordância das respectivas classificações é baixa;
- À semelhança do que foi encontrado para as análises da efectividade e da eficiência, os hospitais apresentam comportamento distinto para os casos cirúrgicos e médicos e por doença. Para esta última perspectiva é igualmente afirmado que, por vezes, existem maiores diferenças intra-hospitalares do que entre hospitais;
- Existem igualmente grandes diferenças nos hospitais na sua efectividade e eficiência, tanto expressa pela diferente classificação dos hospitais, como pelo próprio modelo explicativo do respectivo desempenho. A este propósito, relembre-se o apresentado na discussão de resultados, onde, para a efectividade se observou sempre uma influência da dimensão dos hospitais, medida pela sua produção (assumindo uma distribuição do tipo de U invertido) e para a eficiência, a ausência deste efeito dimensão, para o total das admissões e para os casos médicos e com uma distribuição do tipo U para os casos cirúrgicos.

Neste sentido, em seguida serão discutidos os seguintes aspectos:

- Importância da localização, do tipo e da dimensão do hospital para o respectivo desempenho global;
- Identificação das principais razões justificativas para os diferentes desempenhos entre casos médicos e cirúrgicos;
- Discussão sobre a existência de contradições entre a otimização da efectividade e da eficiência.

Para se investigar a importância da localização, do tipo e da dimensão do hospital serão utilizadas as mesmas variáveis dos estudos da efectividade e da eficiência, sendo igualmente realizadas análises de regressão para os resultados brutos e ajustados pelo risco.

A este propósito, atendendo a que para os resultados brutos, a mortalidade e a demora média têm escalas diferentes, adoptou-se o seguinte procedimento para se apurar o indicador de desempenho global:

Indicador Global Bruto = (Taxa de Mortalidade Corrigida + Demora Média Corrigida) / 2

onde

Taxa de Mortalidade Corrigida do hospital = Taxa de Mortalidade do Hospital / Taxa de Mortalidade Global

e

Demora Média Corrigida do hospital = Demora Média do hospital / Demora Média Global

No Quadro LXXII são apresentados para todas as admissões e para os casos cirúrgicos e médicos, as análises de regressão para os resultados brutos e ajustados pelo risco.

Para os resultados brutos, embora com intensidades diferentes, todas as variáveis independentes (localização, tipo e dimensão do hospital) são explicativas do desempenho dos hospitais, com excepção para os casos cirúrgicos. Apesar desta situação, o nível de explicação global mais elevado é atingido nestes casos.

Para os resultados ajustados pelo risco a situação é completamente diferente:

- Para o total das admissões a única variável explicativa é a localização dos hospitais, com especial relevo para as Regiões Centro e Norte;
- Para os casos cirúrgicos o modelo não revela qualquer explicação, tanto globalmente, como por variáveis;
- Para os casos médicos, todas as variáveis são significativas. Para a produção os hospitais com melhor desempenho são encontrados entre aqueles com maior e menor dimensão (embora com maior relevo no primeiro conjunto de hospitais);

- Para qualquer destes modelos o poder explicativo global do modelo (R^2 ajustado) é sempre menor do que o encontrado para os resultados brutos.

Quadro LXXII
Resultados do Modelo
Desempenho Global: Bruto e Ajustado
Total das Admissões, Casos Cirúrgicos e Casos Médicos

	Todos		Cirúrgicos		Médicos	
	Brutos	Ajustados	Brutos	Ajustados	Brutos	Ajustados
R^2 (ajustado)	0.438	0.200	0.650	0.025	0.531	0.337
Estatística F	7.792*	3.181*	16.768*	1.217	10.848*	5.426*
Constante	Sim	Não	Sim	Não	Sim	Não
Produção (termo simples)	Sim	Não	Não	Não	Sim	Não
Produção (termo ao quadrado)	Sim	Não	Não	Não	Não	Sim
Centrais e Universitários	Sim	Não	Sim	Não	Sim	Sim
H. Especializados	Sim	Não	Sim	Não	Sim	Não
Distritais de Nível 1	Não	Não	Sim	Não	Sim	Não
IPOs	Não	Não	Sim	Não	Não	Não
Alentejo	Não	Não	Não	Não	Não	Não
Algarve	Não	Não	Não	Não	Não	Não
Centro	Sim	Sim	Sim	Não	Sim	Sim
Norte	Sim	Sim	Sim	Não	Sim	Sim

* significativa a 0.05

Estes aspectos comprovam genericamente que a utilização dos resultados brutos para se avaliar o desempenho dos hospitais conduz a conclusões erróneas, pelo que a introdução do ajustamento pelo risco constitui um instrumento necessário para se concretizar esta finalidade.

Por outro lado, para além do diferente comportamento dos hospitais entre os casos médicos e cirúrgicos, constata-se que não existem tipos de hospital, nem dimensões óptimas para se maximizarem os resultados.

Assim, pese embora o facto de as Regiões Centro e Norte apresentarem resultados melhores, deve concluir-se que em Portugal o melhor ou pior desempenho, parece estar mais associado com a gestão dos hospitais do que com características intrínsecas das organizações de saúde.

Contudo, face ao exposto quando se analisa o desempenho por doença para qualquer das dimensões consideradas, no qual avulta a existência de grandes disparidades intra-hospitalares e de diversos hospitais com dos melhores e dos piores desempenhos para cada uma das doenças, parece reforçar-se que o esforço de gestão é mais patente ao nível do serviço do que do próprio hospital.

Ou seja, estes resultados da investigação parecem exprimir que em cada hospital existem vários “hospitais” cada um com a sua dinâmica própria, pelo que se pode ainda concluir que ainda existe uma insuficiência ao nível macro ou estratégico da gestão dos hospitais.

Se tal se deve somente a deficiências das administrações hospitalares, ou deve ser atribuído ao actual enquadramento normativo e modelo de financiamento ou a um “mix” destas situações são questões que não são alvo deste estudo. No entanto, atendendo à importância do fenómeno, bem como ao processo de mudança organizacional em curso, sugere-se que no futuro sejam realizados estudos específicos sobre estes assuntos.

Passando agora a analisar o desempenho global dos hospitais entre os casos cirúrgicos e médicos, conforme foi referido existe um comportamento distinto dos hospitais, bem expresso pelo *K de Cohen* quando se compara a classificação destes (0.035).

Contudo, quando se associa o desempenho dos hospitais medido pelos “z scores”, verifica-se que quanto melhor é o desempenho nos casos médicos melhor é o desempenho nos casos cirúrgicos, visto que o coeficiente de correlação de Pearson é de 0.468 (significativo a 0.01). Conclusão idêntica é retirada quando se associam as ordenações dos hospitais para os casos médicos e cirúrgicos, visto que o coeficiente de correlação de Spearman é de 0.492 (igualmente significativo a 0.01), pelo que existe uma relação directa entre o desempenho nos casos cirúrgicos e médicos.

Assim, pode concluir-se que embora as ordenações sejam diferentes para os casos cirúrgicos e médicos, quanto melhor é o desempenho global nos casos cirúrgicos, melhor é o desempenho global nos casos médicos.

No Quadro LXXIII é apresentada a matriz da ordenação dos hospitais para o desempenho global para os casos cirúrgicos e médicos por quartis.

Quadro LXXIII
Ordenação do Desempenho dos Hospitais por Quartis
Casos Cirúrgicos e Casos Médicos

Casos Cirúrgicos	Casos Médicos			
	Até 1º Quartil	Entre o 1º e o 2º Quartis	Entre o 2º e o 3º Quartis	Depois do 3º Quartil
Até 1º Quartil	11	6	2	2
Entre o 1º e o 2º Quartis	5	8	6	3
Entre o 2º e o 3º Quartis	2	4	10	6
Depois do 3º Quartil	3	4	4	10

Assinalados a azul estão os 39 hospitais que estão classificados no mesmo grupo de desempenho. A vermelho estão identificados 16 hospitais com maiores discrepâncias no desempenho entre os casos cirúrgicos e os médicos. Para os restantes 31 hospitais considera-se que as diferenças não são muito relevantes.

Esta análise vem confirmar as indicações retiradas da análise de correlação, visto que 39 hospitais (45.35%) apresentam o mesmo desempenho para os casos cirúrgicos e médicos, enquanto que somente 18.60.% (16 hospitais) apresentam desempenhos completamente distintos para estes dois grandes conjuntos de doentes.

Para os 16 hospitais com desempenhos mais distintos entre os casos cirúrgicos e médicos, 8 são do Tipo III, 3 dos Tipos I e IV e 2 do Tipo V, estando 6 localizados nas Regiões IV e V e 4 na III.

Estes valores parecem apontar para a inexistência de um padrão justificativo do comportamento dos hospitais para estes casos. Contudo, deve referir-se que tanto nos hospitais do Tipo II, como nas Regiões de Saúde I e II o comportamento destas organizações de saúde é semelhante entre os casos cirúrgicos e médicos.

No que se refere ao desempenho dos hospitais para a efectividade e para a eficiência, tanto na apresentação dos resultados, como na discussão sobre estas dimensões, foi referida a existência de indicações contrárias e potencialmente contraditórias.

Na realidade, o *K de Cohen* é de 0.000, de 0.035 e de 0.011 quando se comparam as classificações dos hospitais na efectividade e na eficiência, para o total das admissões, para os casos cirúrgicos e para os casos médicos, respectivamente.

O valor desta estatística para as três situações analisadas permite concluir pela existência de um comportamento completamente diferente dos hospitais na efectividade e na eficiência. A este propósito deve referir-se que resultados semelhantes foram encontrados por Zimmerman e outros (1993), quando comparam a efectividade e a eficiência entre hospitais universitários e hospitais não universitários.

Por outro lado, os valores do coeficiente de correlação de Spearman entre as ordenações dos hospitais para a efectividade e a eficiência são de 0.113, de 0.114 e de 0.106 (todos não significativos), respectivamente para o total das admissões, casos cirúrgicos e médicos, parecendo igualmente traduzir a inexistência de uma associação no desempenho dos hospitais para a efectividade e para a eficiência.

Os coeficientes de correlação de Pearson entre os “z scores” da efectividade e da eficiência é de 0.049, de 0.032 e de 0.113 (todos não significativos), respectivamente para o total das admissões, casos cirúrgicos e médicos. Pelo que mais uma vez se pode concluir pela inexistência de uma associação entre estas duas variáveis.

No entanto, atendendo por um lado, a que todas as associações são positivas e a que, por outro lado, nunca se observam correlações significativas e com sinal negativo, pode concluir-se que embora o desempenho dos hospitais seja diferente para a efectividade e para a eficiência, não existem contradições entre estas duas dimensões.

Ou seja, para se otimizar a efectividade não é necessário sacrificar a eficiência, sendo o inverso igualmente verdadeiro. Bons exemplos ilustrativos desta situação são encontrados nos hospitais 13, 32, 41 e 48 os quais ocupam sempre as 10 primeiras posições para a efectividade, para a eficiência e para o desempenho global.

Os hospitais 12, 29, 40 e 43, embora por motivos diferentes, ilustram igualmente esta situação, visto que ocupam sempre as 10 últimas posições para a efectividade, para a eficiência e para o desempenho global.

No Quadro LXXIV é apresentada a matriz da ordenação dos hospitais para a efectividade e para a eficiência por quartis.

Quadro LXXIV
Ordenação do Desempenho dos Hospitais por Quartis
Efectividade e Eficiência

	Eficiência			
Efectividade	Até 1º Quartil	Entre o 1º e o 2º Quartis	Entre o 2º e o 3º Quartis	Depois do 3º Quartil
Até 1º Quartil	8	6	2	6
Entre o 1º e o 2º Quartis	5	8	5	4
Entre o 2º e o 3º Quartis	3	5	9	5
Depois do 3º Quartil	6	3	6	7

Assinalados a azul estão os 32 hospitais que estão classificados no mesmo grupo de desempenho. A vermelho estão identificados 24 hospitais com maiores discrepâncias no desempenho entre a efectividade e a eficiência. Para os restantes 32 hospitais considera-se que as diferenças não são muito relevantes.

Para os 24 hospitais com desempenhos mais distintos entre a efectividade e a eficiência deve referir-se que 12 correspondem a situações com efectividade bastante superior à eficiência, ocorrendo nos restantes 12 exactamente o contrário. Destes hospitais, 15 são do Tipo III, 6 do Tipo I e 1 dos Tipos II, IV e V, estando 12 localizados na Região IV, 6 na Região III, 5 na IV e 1 na II.

Para além das frequências absolutas elevadas encontradas nos hospitais do Tipo III (15) e para a situação da Região de Saúde IV (12), existem outros indícios que sugerem a existência de um padrão para o diferente desempenho entre a efectividade e a eficiência.

De facto observa-se que para os 15 hospitais do Tipo III com grandes discrepâncias, 9 correspondem a situações em que a eficiência é melhor que a efectividade. Nos 12 hospitais da Região IV 10 apresentam igualmente níveis de eficiência superiores aos da efectividade. Na situação contrária encontram-se os hospitais localizados na Região de Saúde III, visto que todos apresentam níveis de efectividade superiores aos da eficiência.

No entanto, bem mais relevante é o facto de os 9 hospitais do Tipo III que apresentam melhor eficiência estarem todos localizados na Região de Saúde IV, o que parece evidenciar que as razões para as grandes discrepâncias são encontradas nesta Região e ainda na Região III, embora neste particular com sentido contrário, visto que a efectividade é melhor do que a eficiência.

Para melhor visualizar a dimensão deste fenómeno são apresentados os coeficientes de correlação entre os “z scores” da efectividade e da eficiência para todas as admissões e para todos os hospitais (Cenário A), para os 64 hospitais com nenhuma ou pequenas discrepâncias entre a efectividade e a eficiência (Cenário B), para os 76 hospitais em que estão excluídos os 12 com níveis de eficiência bastante superiores aos da efectividade (Cenário C) e para os 76 hospitais em que estão excluídos os 12 com níveis de efectividade bastante superiores aos da eficiência (Cenário D) (ver Quadro LXXV).

Quadro LXXV
Coeficientes de Correlação entre a
Efectividade e a Eficiência dos Hospitais

	Eficiência			
	Cenário A	Cenário B	Cenário C	Cenário D
Efectividade	0.049	0.632 *	0.196	0.396 *

* significativo a 0.01

Estes resultados provam que, embora para o total dos hospitais e para o conjunto dos hospitais onde não estão incluídos os 12 com maior eficiência em relação à efectividade (Cenário C) não exista uma correlação significativa. Para os restantes cenários, os que excluem os 24 hospitais com grandes

discrepâncias entre estas duas dimensões (cenário B) e os que excluem os 12 hospitais com maior efectividade em relação à eficiência (Cenário D), existem correlações significativas e de sinal positivo. Ou seja, para estes dois últimos cenários quanto maior é a efectividade, maior é a eficiência.

Assim, pode concluir-se o seguinte:

- Para todas as admissões não existe qualquer associação entre a efectividade e a eficiência;
- São identificados 24 hospitais em que as discrepâncias no desempenho destas duas dimensões são bastante relevantes, sendo 12 com maior efectividade e os restantes 12 com maior eficiência;
- Para a situação em que os hospitais apresentam maior efectividade, 50% das organizações de saúde estão localizadas na Região de Saúde III;
- Na situação contrária, hospitais com maior eficiência, 10 hospitais estão localizados na Região de Saúde IV, ainda com a particularidade de 9 corresponderem a hospitais do Tipo III;
- A não associação entre a efectividade e a eficiência não corresponde a um fenómeno estrutural em Portugal, visto que com a exclusão dos 24 hospitais com grandes discrepâncias entre estas duas dimensões ou somente dos 12 hospitais com maiores níveis de efectividade, a correlação é directa e significativa.

6.3. Implicações da Avaliação da Produção e do Desempenho Hospitalar no Sector de Saúde

Em função dos resultados do estudo, bem como de aspectos evidenciados durante a discussão, interessa agora discutir algumas das implicações deste modelo de avaliação no sistema da saúde.

Estas implicações serão analisadas em relação a três aspectos estruturantes do Serviço Nacional de Saúde – o financiamento, a regulação e a gestão dos hospitais, bem como no que se refere à publicitação de resultados de avaliação do desempenho dos hospitais.

Em relação ao financiamento, vai-se impor desde já uma restrição, somente será abordada uma modalidade de pagamento prospectivo – o pagamento prospectivo por caso – essencialmente porque é o modelo que teoricamente está implementado em Portugal.

Neste sentido, as implicações deste estudo, devem ser consideradas, tanto a montante, ou seja no mecanismo de definição dos preços, como a jusante, ou seja, na definição de um sistema de incentivos que premeie o desempenho.

Para a definição de preços convém recordar que a mesma é feita de acordo com os DRGs, traduzindo a complexidade dos casos tratados, sendo publicada no Diário da República, tendo a referida tabela sido periodicamente revista e actualizada desde que está implementado este sistema de definição de preços hospitalares.

Para além de diversos problemas conceptuais e operacionais identificados noutros estudos (Costa, 1990; Miguel e Costa, 1997 e Santana, 2003) e que não serão objecto de análise nesta discussão, interessa agora evidenciar alguns dos resultados apresentados e discutidos:

- Atendendo a que os DRGs e o Disease Staging estão a medir dimensões diferentes, a complexidade e a gravidade, respectivamente, a produção dos hospitais, medida pelos respectivos Índices de Casemix, é completamente diferente;
- O perfil das admissões dos hospitais, igualmente medido pela complexidade e pela gravidade é igualmente diferente, configurando-se os hospitais de maior dimensão como aqueles que tratam doentes mais complexos e os hospitais de menor dimensão os que tratam doentes mais graves. Por sua vez, são igualmente os hospitais com maior dimensão, aqueles que maximizam a diferença entre a complexidade e a gravidade dos casos tratados.

Por outro lado, alguns dos comportamentos esperados do pagamento prospectivo por caso (por DRG) respeitam à preferência por casos cirúrgicos

(Coffey e Goldfarb, 1986), à escolha de doentes menos graves por DRG (Rosko, 1988) e à criação de eventuais conflitos com a qualidade dos cuidados prestados (Aronow, 1988).

Assim, a conjugação dos aspectos anteriormente referidos permite as seguintes sugestões:

- A definição de um mecanismo de preços com base em custos por doença (Santana, 2003), sendo para tal necessário que se desenvolva um sistema de custeio por doente;
- A introdução de níveis de gravidade para diferenciar preços dentro do mesmo produto. Para tal podem ser utilizadas as indicações do Disease Staging, ou de uma das versões dos DRGs que incluem a gravidade (APRDRGs ou IRDRGs) ou ainda de um sistema que utilize a informação clínica;
- Definição da demora média central por DRG não em função do que foi feito, mas da demora média esperada. Mais uma vez a utilização dos sistemas administrativos ou clínicos podem introduzir informações válidas para o efeito.

No que se refere à definição do sistema de incentivos, dois tipos de questões devem ser abordadas: (1) o sistema de financiamento deve ou não ser replicado aos profissionais de saúde; (2) qual o melhor mecanismo para premiar ou punir os hospitais e/ou prestadores?

Em relação ao primeiro aspecto, generalização do sistema de financiamento, deve desde já referir-se que faz muito pouco sentido que um sistema de pagamento pela produção não tenha repercussão nos profissionais de saúde.

Em Portugal, esta afirmação ganha ainda mais relevo, visto que até ao presente momento não existe liberdade de escolha por parte dos consumidores. Nesta conformidade, penalizar hospitais com menos produção, somente compromete a qualidade dos cuidados prestados, visto que o pagamento aos profissionais de saúde assume um carácter de quase despesas fixas. Na situação contrária, o anacronismo está igualmente presente, visto que não se encontram argumentos que justifiquem um maior pagamento a hospitais com melhores níveis produtivos (afinal nos hospitais onde potencialmente existem menos problemas com a procura) sem repercussões nos pagamentos aos profissionais (qual é o destino do excedente financeiro?).

Neste sentido, somente será possível concluir que o pagamento prospectivo por caso ganha mais sentido quando existir a possibilidade de ser replicado aos profissionais de saúde. Se os mecanismos são definidos ao nível central ou por hospital é uma questão que escapa ao âmbito deste estudo.

No que se refere ao esquema de incentivos e aceite o princípio da sua replicabilidade aos prestadores, deve igualmente acentuar-se que existe uma necessidade de mudança de paradigma.

Assim, enquanto que a discussão actualmente presente em Portugal acentua a necessidade de se diferenciarem preços por tipo do hospital, sugere-se que na eventualidade de se criarem preços diferenciais, os mesmos sejam fixados em função de níveis de desempenho.

Para tal podem, por exemplo, ser criados grupos de hospitais em função do desempenho, pagando marginalmente acima para as organizações com melhor desempenho e marginalmente abaixo para as que apresentam pior desempenho.

Naturalmente, nesta perspectiva ganha relevo a escolha e a valorização das dimensões do desempenho – efectividade? eficiência? ambas? Outras? Quais?

Se a opção for por um esquema misto deve ainda discutir-se qual a ponderação que deve ser atribuída a cada uma das dimensões.

Finalmente, deve reiterar-se que este esquema de incentivos somente ganha relevo se existirem preços diferenciados por produto em função da gravidade e que este mecanismo deve ser replicado sobre os profissionais de saúde.

No que se refere à regulação, para além dos aspectos referidos anteriormente, outras duas perspectivas são importantes de discutir:

- O perfil das admissões e a mobilidade dos doentes;
- As divergências no desempenho dos hospitais.

Em relação ao perfil das admissões recordem-se alguns dos aspectos mais salientes:

- Existem diferenças no comportamento dos hospitais no que se refere à complexidade dos casos tratados, as quais por sua vez são mais intensas nos casos mais complexos (aspecto presente para todas as admissões e para os casos cirúrgicos e médicos);
- Para a gravidade dos casos a situação é praticamente idêntica à referida para a complexidade, ou seja maiores diferenças entre hospitais nos casos mais graves, embora com uma maior atenuação nos casos médicos;
- A principal diferença entre estas duas perspectivas é encontrada nos hospitais com maiores e menores complexidade e gravidade, visto que são os hospitais do Tipo I os que tratam doentes mais complexos,

enquanto que para os casos mais graves tal ocorre nos hospitais do Tipo IV.

Este último aspecto parece indiciar a existência de informações e de comportamentos distintos por parte dos hospitais para a complexidade e para a gravidade dos casos tratados. Na realidade este fenómeno foi confirmado, essencialmente nos seguintes aspectos:

- Para todas as admissões, os hospitais com maiores diferenças entre a complexidade e a gravidade dos casos são do Tipo I e V, enquanto que com menores diferenças são identificados os hospitais do Tipo IV e III. Observou-se ainda que não existe associação entre a complexidade e a gravidade dos casos tratados e que a gravidade varia inversamente com o aumento da produção, enquanto que a complexidade e a diferença entre a complexidade e a gravidade dos casos varia directamente com o aumento da produção.

Constatou-se ainda que este fenómeno não é estrutural no panorama hospitalar português, visto que uma análise por doença permitiu concluir por uma relação directa entre a complexidade e a gravidade dos casos tratados.

Em relação à mobilidade dos doentes, aspecto ainda não directamente tratado neste estudo, procedeu-se a uma análise tendo em vista os seguintes aspectos:

- Considerou-se o distrito como variável indicadora da localização dos hospitais;
- Considerou-se igualmente o distrito como variável indicadora da origem dos doentes;
- Construiu-se uma variável dicotómica para cruzar a localização dos hospitais com a origem dos doentes (*Localorig*) com seguintes valores:
 - *Localorig* = 0, quando os distritos de prestação e de origem dos doentes são os mesmos;
 - *Localorig* = 1, quando os distritos de prestação e de origem dos doentes são diferentes.
- Identifica-se a percentagem de doentes por hospital para a variável *Localorig*;
- Identifica-se a complexidade e a gravidade dos doentes por *localorig*. Para a complexidade utiliza-se o Índice de Casemix dos DRGs e para a gravidade o Índice de Casemix do Disease Staging.

No Quadro LXXVI são apresentados por Tipo de hospital a percentagem de episódios em que os distritos de origem dos doentes e de prestação de cuidados coincidem (*Localorig*=0) e a percentagem de doentes em que os distritos de origem dos doentes e de prestação de cuidados não coincidem (*Localorig*=1),

bem como a complexidade e a gravidade dos casos tratados para estas duas variáveis.

Quadro LXXVI
Origem dos Doentes e Prestação de Cuidados
Complexidade e Gravidade dos Casos

Hospitais	Localorig = 0 (%)	Localorig = 1 (%)	Localorig = 0 Complexidade	Localorig = 1 Complexidade	Localorig = 0 Gravidade	Localorig = 1 Gravidade
Tipo I	74.71	25.29	1.26	1.55	1.20	0.87
Tipo II	81.60	18.40	0.76	1.02	0.06	0.09
Tipo III	94.41	5.59	0.98	0.98	0.97	0.66
Tipo IV	94.72	5.28	0.94	0.95	1.20	0.70
Tipo V	47.95	52.05	1.84	1.87	1.52	1.13
Total	85.64	14.36	1.07	1.43	1.03	0.81

Para a globalidade dos hospitais existem 14.4% dos doentes que são tratados fora do seu local de origem, com uma maior expressão nos hospitais do Tipo V (52%) e do Tipo I (25.3%) e com menor expressão nos hospitais do Tipo IV (5.28%) e do Tipo III (5.59%).

Na globalidade os doentes tratados fora do seu local de origem são mais complexos e menos graves. Para a complexidade é nos hospitais do Tipo I e II, onde este fenómeno está mais patente, visto que para os restantes Tipo de hospitais as diferenças são mais reduzidas. Para a gravidade, somente nos hospitais do Tipo II é observado um comportamento diferente, dado que a gravidade dos doentes com origem distinta do local de prestação é maior que a dos restantes casos. Por sua vez, os hospitais do Tipo IV e III são os que apresentam maiores diferenças entre a gravidade dos doentes tratados com origem igual ao local de prestação, quando comparados com os restantes doentes.

Assim, pode concluir-se que na sua globalidade os doentes com maior mobilidade (sendo esta medida pela não coincidência entre o distrito de origem dos doentes e o distrito de prestação de cuidados) são simultaneamente mais complexos e menos graves.

Neste sentido, para a política/perfil das admissões hospitalares, a função regulação deve ser exercida nas seguintes vertentes:

- A conformidade destes comportamentos com os princípios e finalidades do Serviço Nacional de Saúde Português;
- A definição de orientações sobre a referência de doentes, tanto a montante, como a jusante;
- A definição de mecanismos e de indicadores para monitorizar a mobilidade dos doentes;

- A definição de uma política de incentivos (incluído penalizações) para premiar ou punir os hospitais e/ou prestadores que apresentem um comportamento não esperado.

No que se refere ao desempenho dos hospitais foi referenciado ao longo do estudo a existência de discrepâncias significativas de hospitais do mesmo Tipo e de hospitais dentro da mesma Região de Saúde. Estas diferenças são encontradas, tanto na análise da efectividade dos hospitais, como na eficiência dos hospitais, ou ainda nas diferenças entre a efectividade e a eficiência.

Para a efectividade e por Tipo de hospital podem evidenciar-se as seguintes situações:

- Existem 44 hospitais com piores resultados que os esperados. Nos hospitais do Tipo V observa-se que 66.7% destas organizações estão nesta situação (hospitais 87 e 88), nos hospitais do Tipo IV este valor é de 65.2% (merecendo os hospitais 66, 72 e 81 particular destaque), nos hospitais do Tipo III 52.5% apresentam resultados piores que os esperados (com especial relevo para os hospitais 23, 29, 30, 33, 40, 43, 50, 55, 56, 57, 59 e 62) e nos hospitais do Tipo I este valor é de 37.5% (hospitais 1, 12 e 16, com especial destaque).

Para a efectividade e por Região de Saúde podem referir-se os seguintes aspectos:

- A Região de Saúde III é a que apresenta maior quantidade de hospitais com melhor desempenho que o esperado (65%), imediatamente seguida pela Região de Saúde V (59%). As restantes regiões apresentam números superiores de hospitais com desempenho pior que o esperado, de 66.7% nas Regiões II e IV e de 80% na Região de Saúde I;
- Apesar desta situação é nas Regiões de Saúde III e V que se encontram maiores variações no desempenho entre os hospitais, sendo a maior homogeneidade encontrada na Região I, embora neste particular por maus motivos, visto que a grande maioria dos hospitais apresenta um desempenho pior do que o esperado.

No que se refere à eficiência e por Tipo de hospital a situação é a seguinte:

- Nos hospitais do Tipo V todos as organizações apresentam resultados piores que os esperados (o hospital 88 merece especial destaque), nos hospitais do Tipo II este valor é de 66.7% (merecendo o hospital 19 particular destaque), nos hospitais do Tipo IV 60.9% apresentam resultados piores que os esperados (com especial relevo para o hospital 75), nos hospitais do Tipo III este valor é de 47.5% (com especial destaque nos hospitais 27, 29, 36, 38, 40, 43, 44, 52, 54, 60 e 61) e nos hospitais do Tipo I este valor é de 31.3% (hospitais 4, 9, 10, 12 e 14, com especial destaque).

Para a eficiência e por Região de Saúde podem referir-se os seguintes aspectos:

- A Região de Saúde V é a que apresenta maior quantidade de hospitais com melhor desempenho que o esperado (59%), imediatamente seguida pela Região de Saúde IV (52%). As restantes regiões apresentam valores superiores de hospitais com desempenho pior que o esperado, de 53.9% na Região III, 80% na Região I e de 100% na Região de Saúde II;
- Apesar desta situação é nas Regiões de Saúde III, IV e V que se encontram maiores variações no desempenho entre os hospitais, sendo a maior homogeneidade encontrada na Região II, embora neste particular por maus motivos, visto que todos os hospitais apresentam um desempenho pior do que o esperado.

Para o desempenho global e por Tipo de hospital, observa-se o seguinte:

- Nos hospitais do Tipo V 66.7% das organizações apresentam resultados piores que os esperados (o hospital 88 merece especial destaque), nos hospitais do Tipo IV este valor é de 65.2% (merecendo os hospitais 72 e 75 particular destaque), nos hospitais do Tipo III 65% apresentam resultados piores que os esperados (com especial relevo para os hospitais 23, 29, 30, 33, 38, 40, 43, 52, 55, 56, 59, 60 e 62), nos hospitais do Tipo II este valor é de 50% e nos hospitais do Tipo I este valor é de 31.3% (hospitais 12 e 14, com especial destaque).

Para o desempenho global e por Região de Saúde podem referir-se os seguintes aspectos:

- A Região de Saúde III é a que apresenta maior quantidade de hospitais com melhor desempenho que o esperado (54%), imediatamente seguida pela Região de Saúde V (52%). As restantes regiões apresentam números superiores de hospitais com desempenho pior que o esperado, de 70% na Região IV e de 100% nas Regiões I e II;
- Apesar desta situação é nas Regiões de Saúde III, IV e V que se encontram maiores variações no desempenho entre os hospitais, sendo a maior homogeneidade encontrada nas Regiões I e II, embora neste particular por maus motivos, visto que todos os hospitais apresentam um desempenho pior do que o esperado.

Para as diferenças no desempenho entre a efectividade e a eficiência pode observar-se o seguinte:

- Nos 20 hospitais com maiores diferenças 13 são do Tipo III e 5 do Tipo I. Os restantes 2 hospitais são dos Tipo IV e V. Destes 8 estão localizados na Região de Saúde IV, 6 na Região V, 5 na Região III e 1 na Região II;

- Para os 20 hospitais com menores diferenças 10 são do Tipo III, 4 do Tipo e 6 do Tipo V. Destes 7 estão localizados na Região de Saúde III, 6 na Região V, 6 na Região IV, 1 na Região I e 1 na Região II.

Estes valores exprimem ainda com maior intensidade as diferenças no desempenho dos hospitais, tanto por Tipo, como por localização, quando esta é medida pela Região de Saúde.

Neste sentido, sugere-se o seguinte:

- A realização de estudos mais detalhados para se investigarem as causas para as diferenças apontadas;
- A definição de políticas e de medidas que contribuam para a redução nas disparidades no desempenho hospitalar;
- A criação de um sistema de informação para monitorizar o desempenho dos hospitais.

No que se refere à gestão dos hospitais as questões que se colocam são exactamente as mesmas que se referiram para a função regulação, mas neste particular com especial incidência nas questões internas de cada hospital e ainda tendo em função o comparador para o melhor desempenho (“benchmarking”).

Todas as questões referentes ao perfil das admissões e ao desempenho dos hospitais foram apresentadas e discutidas anteriormente, pelo que somente se reitera a existência de admissões distintas dentro do mesmo hospital, para casos médicos e cirúrgicos e por doença e que, por exemplo, para os 88 hospitais analisados, 37 apresentam grandes discrepâncias no desempenho, quando estas são medidas pela presença entre os melhores e piores, para todos os episódios de internamento, para os casos médicos e cirúrgicos e por doença.

Atendendo a que a dimensão deste fenómeno é ainda maior, bastando para tal não considerar como grandes discrepâncias a presença entre os cinco melhores ou piores hospitais, sugere-se que os procedimentos referidos anteriormente para a regulação sejam replicados internamente, tanto por serviço, como por médico responsável pelo tratamento.

Finalmente, serão discutidos os aspectos relacionados com a publicitação de resultados de avaliação do desempenho dos hospitais.

Deve desde já esclarecer-se que se defende que os resultados da produção e do desempenho dos hospitais devem ser do conhecimento público.

Esta afirmação ganha ainda mais consistência na eventualidade de se pretender eleger a liberdade de escolha como uma mecanismo de política de saúde em Portugal.

Contudo, atendendo a que a liberdade de escolha sugere ainda a necessidade de se considerarem outros elementos e mecanismos, passa-se de seguida a discutir a publicitação dos resultados da actividade dos hospitais no actual modelo organizacional.

Este modelo, como é do conhecimento geral, parte da definição de uma área de atracção dos hospitais a qual inclui igualmente uma rede de referência dos cuidados de saúde.

É neste contexto que existem diversos argumentos que defendem que a publicitação dos resultados dos hospitais só servirá para “penalizar” os consumidores, visto que não dispõem de quaisquer instrumentos para alterar a situação.

No entanto, em relação a estes argumentos, deve ter-se em atenção que a publicitação dos resultados dos hospitais permite, no actual contexto, evidenciar as seguintes vantagens: reduz a assimetria de informação existente no sector da saúde; permite que a actividade dos hospitais seja mais transparente, em detrimento do actual modelo “fechado” existente e permite ainda que as imagens sobre o funcionamento dos hospitais que actualmente são fraccionadas ou esporádicas, muitas vezes decorrentes de opiniões explicitadas pela comunicação social, passem a ter um cenário mais sistematizado e científico (naquilo que se pode designar por imagens fotográficas para filmes).

Face ao exposto, facilmente se compreende a afirmação feita anteriormente de que a publicitação dos resultados sobre a avaliação da produção e do desempenho dos hospitais deve constituir um imperativo do Serviço Nacional de Saúde e que, eventualmente poderá ainda contribuir para as mudanças e melhorias que se consideram necessárias em Portugal.

Como foi igualmente referido, esta publicitação ganha ainda mais relevo com uma possível implementação da liberdade de escolha por parte dos consumidores, visto que constitui um instrumento necessário e válido para fundamentar esta decisão.

Contudo, qual o âmbito desta liberdade de escolha? Por hospital? Por prestador?

Quando se refere a dimensão hospital está subjacente que um consumidor escolhe uma determinada organização de saúde, a qual é posteriormente responsável pela sua situação de saúde, sejam os cuidados prestados na mesma ou em qualquer outra organização.

As principais vantagens desta abordagem consistem na maior facilidade de implementação e de gestão e inclusivamente na definição de mecanismos de financiamento. Os principais inconvenientes estão associados ao facto, de existirem desempenhos distintos dentro de uma organização de saúde (esta questão não é meramente académica, como os resultados deste estudo provam), pelo que o consumidor poderá ter sempre presente que em determinada situação não está a ter acesso aos melhores cuidados (independentemente da metodologia utilizada para a medição do desempenho).

Por outro lado, a escolha do prestador, que nesta linha de pensamento está associada ao princípio de o consumidor poder escolher casuisticamente a organização responsável pela prestação de cuidados, embora apresente como principal vantagem o inconveniente referido para a alternativa anterior – acessibilidade aos melhores cuidados, apresenta alguns problemas que convém discutir.

Desde logo, a questão do financiamento. Embora se possa argumentar que um modelo de pagamento por caso possa obviar a alguns problemas, também não é menos verdade que pode originar uma multiplicação de contactos e de actos, o que para além de aumentar as despesas, pode ainda introduzir ineficiência no sistema.

Outra questão respeita à circulação da informação. De facto, uma escolha casuística do prestador, para não originar problemas suplementares na efectividade e na eficiência dos cuidados prestados, apresenta maiores exigências ao sistema de informação.

Exigências de nível conceptual, deontológico e ético que passam pela consideração que toda a informação clínica é do doente e não do prestador, pelo que deve ser do conhecimento de todos os potenciais prestadores, desde que o próprio dê a devida autorização.

Exigências igualmente ao nível operacional, essencialmente na arquitectura do sistema de informação, visto que para esta circular é necessária uma maior automatização e compatibilização dos suportes e “outputs” de informação.

Como é óbvio, neste estudo não se pretende, nem esgotar o assunto, nem dar respostas sobre a melhor alternativa para se implementar um mecanismo de liberdade de escolha.

No entanto, sugere-se que todas as questões referidas que vão desde o financiamento, passando pela regulação e pela gestão das organizações de saúde, bem como a publicitação de resultados, sejam devidamente equacionadas antes de implementadas e que, mesmo um conceito aparentemente tão simples como a liberdade de escolha, seja igualmente reflectida antes de se aplicar o modelo que se pretende.

CONCLUSÕES

7. Conclusões

O presente estudo apresentava como principais objectivos a definição e utilização de dimensões para acompanhar a produção dos hospitais e a operacionalização de um modelo de avaliação do desempenho hospitalar.

Paralelamente são ainda discutidas algumas das razões justificativas para os distintos desempenhos dos hospitais e algumas das principais implicações para a definição de políticas de saúde decorrentes da utilização destes mecanismos e instrumentos.

Para a medição da produção foram consideradas duas perspectivas: a medição da produção através dos Índices de Casemix e o perfil das admissões.

No que se refere à medição da produção, a principal originalidade consiste na introdução da gravidade para medir o “casemix”, tendo-se chegado às seguintes conclusões:

- A produção total do hospital é valorizada de forma diferente em função da utilização dos DRGs ou do Disease Staging;
- Os DRGs valorizam mais os casos cirúrgicos, enquanto que o Disease Staging valoriza mais os casos médicos;
- Os DRGs e o Disease Staging atribuem igualmente ponderações distintas em função das doenças;
- Os DRGs facultam diferenças menores na ponderação da produção, tanto por gravidade dos doentes, como por idade.

Tentando identificar as razões que explicam as diferenças referidas entre estes dois tipos de Índice de Casemix, foi evidenciado que o tipo de tratamento parece ser o factor mais decisivo para o desempenho distinto dos dois sistemas de classificação de doentes.

Este aspecto pode estar associado à própria lógica de construção dos dois sistemas de classificação de doentes em análise. De facto, enquanto que os DRGs se baseiam na identificação de produtos com semelhanças no consumo de recursos e cuja principal finalidade é o financiamento dos hospitais, o Disease Staging centra a sua lógica de desenvolvimento no modelo biológico das doenças, com as naturais vantagens em termos de avaliação da qualidade dos cuidados prestados.

Neste sentido foi construído um Índice de Casemix Composto (média não ponderada dos Índices de Casemix dos DRGs e do Disease Staging), tendo-se concluído que a metodologia do Índice Composto parece a mais adequada, visto que incorpora simultaneamente e com o mesmo peso relativo informação sobre a complexidade e sobre a gravidade dos casos tratados em cada hospital.

No entanto, atendendo à importância do assunto sugeriu-se ainda o seguinte:

- A disponibilização aos hospitais dos três indicadores;
- A escolha dos melhores sistemas de classificação de doentes para serem aplicados aos hospitais;
- A investigação sobre a melhor metodologia para se atribuírem ponderações para cálculo do Índice de Casemix Composto.

Para se qualificar o perfil das admissões hospitalares considerou-se a concentração / diversificação da produção e a complexidade e a gravidade dos casos tratados.

No que se refere à caracterização da concentração da produção foram utilizados dois indicadores, um referente à raridade dos casos em cada hospital (percentagem de produtos diferentes por hospital) e o segundo respeitante ao nível de concentração da produção por hospital.

No que se refere ao efeito raridade observou-se que, pese embora o facto de o Disease Staging apresentar um número de produtos diferentes por hospital mais elevado que o proporcionado pelos DRGs, os dois sistemas de classificação de doentes estão genericamente a disponibilizar a mesma informação para qualificar o efeito raridade na produção do hospital.

Observou-se igualmente que existe um comportamento heterogéneo entre os hospitais, perspectivando-se a dimensão dos hospitais, medida pela quantidade produzida, como um dos factores que mais contribuem para as diferenças detectadas. Inclusivamente, foi referido que o efeito raridade da produção varia proporcionalmente com o aumento da produção.

Quando se analisa a concentração da produção, medida pelo número de produtos diferentes responsáveis por 80% da produção em cada hospital observou-se o seguinte:

- Os DRGs e o Disease Staging estão genericamente a dar a mesma informação sobre a concentração da produção nos hospitais, embora este último sistema de classificação de doentes apresente níveis de concentração mais elevados;
- A existência de um comportamento heterogéneo entre hospitais, mas mais atenuado do que o presente no efeito raridade;
- Importância do volume de produção sobre a respectiva concentração, está presente com os DRGs e com o Disease Staging, embora de uma forma mais intensa no primeiro sistema de classificação.

Para a complexidade e a gravidade dos casos tratados, utilizou-se a informação disponibilizada pelos DRGs para se caracterizar o primeiro atributo e pelo Disease Staging para se apurar a gravidade dos doentes.

Em termos gerais constatou-se que a complexidade e a gravidade dos casos disponibilizam informações diferentes sobre a actividade dos hospitais, sendo os casos cirúrgicos mais complexos e os casos médicos mais graves.

Uma análise mais desagregada identificou ainda que existem comportamentos distintos por parte dos hospitais para a complexidade e para a gravidade dos casos tratados. Na realidade este fenómeno foi confirmado, essencialmente no seguinte:

- Para todas as admissões existem diferenças entre a complexidade e a gravidade dos casos tratados em função do Tipo de hospital;
- Observou-se ainda que não existe associação entre a complexidade e a gravidade dos casos tratados e que a gravidade varia inversamente com o aumento da produção, enquanto que a complexidade e a diferença entre a complexidade e a gravidade dos casos varia directamente com o aumento da produção;

Atendendo a que este aspecto não é estrutural no panorama hospitalar português, visto que uma análise por doença permitiu concluir por uma relação directa entre a complexidade e a gravidade dos casos tratados, apuraram-se de seguida algumas das vantagens dos hospitais derivadas do seu perfil de admissões.

Os principais resultados foram os seguintes:

- A taxa de mortalidade e a demora média variam directamente com a complexidade e com a gravidade dos doentes tratados por hospital, tanto considerados isoladamente, como conjuntamente;
- Observou-se igualmente que quanto maior a diferença entre a complexidade e a gravidade dos casos tratados menores são a taxa de mortalidade e a demora média.

Estes resultados evidenciam uma enorme variedade de situações anómalas ou, no mínimo, não esperadas.

Até ao presente momento, a informação existente em Portugal sobre a actividade dos hospitais utilizava a informação constante dos resumos de alta dos hospitais e com recurso a um sistema de classificação de doentes – os Diagnosis Related Groups.

Em função desta informação os hospitais apresentavam um comportamento heterogéneo em relação ao seu perfil de admissões.

No entanto, este resultado aparentemente não conflituava com a universalidade do Serviço Nacional de Saúde, visto que os hospitais com maior Índice de Casemix e que naturalmente tratavam casos mais complexos eram simultaneamente aqueles que produziam mais.

Atentas igualmente as especificidades da organização da oferta em Portugal é natural deduzir que a hospitais com maior produção correspondem organizações de saúde com maior diferenciação, normalmente qualificados como hospitais universitários, centrais e especializados, como é o caso dos Institutos de Oncologia.

Ainda dentro desta linha de pensamento deve igualmente referir-se que os hospitais que tratavam casos menos complexos apresentavam resultados brutos, nomeadamente a taxa de mortalidade e a demora média, melhores que os outros hospitais, os quais tratavam doentes mais complexos.

Estas premissas e resultados eram frequentemente utilizados como justificativos dos diferentes desempenhos dos hospitais, argumentando-se ainda frequentemente que os hospitais de maior dimensão produtiva, por estarem no fim da linha hierárquica de referência, seriam os que teoricamente teriam menos possibilidade de escolher doentes e concomitantemente tratariam os doentes mais graves.

As consequências naturais deste fenómeno eram de dupla natureza. Em primeiro lugar, é muito difícil comparar a actividade dos hospitais, visto que os mesmos tratam doentes completamente diferentes e, em segundo lugar, a maior complexidade e gravidade dos casos tratados era, só por si, justificativa de uma maior necessidade de recursos humanos, financeiros e tecnológicos para os hospitais universitários, centrais e especializados.

A utilização da severidade do estado do doente, mediante o recurso ao Disease Staging, vem alterar substancialmente as conclusões atrás referidas.

De facto, embora se mantenha a relação directa entre a gravidade e os resultados dos hospitais, ou seja quanto maior a gravidade dos doentes maior a taxa de mortalidade e a demora média, os hospitais que tratam doentes mais graves são os de menor dimensão produtiva, afinal aqueles que na lógica organizativa portuguesa correspondem maioritariamente a hospitais distritais de nível 1 (19 hospitais num total de 22).

Esta situação constitui desde já uma novidade no actual panorama hospitalar português, dado que se esperava que os hospitais com maior dimensão

tratassem os doentes mais graves, à semelhança do que se verifica com a complexidade dos casos.

Atendendo a que estes valores não parecem corresponder a nenhum fenómeno estrutural, visto que uma análise por doença vem comprovar uma associação directa e significativa entre a complexidade e a gravidade dos casos tratados, o comportamento distinto dos hospitais deve ser procurado noutras razões.

Para tal as principais hipóteses de investigação parecem apontar para questões associadas com a lógica organizacional dos hospitais portugueses, com a natureza específica das doenças e da respectiva gravidade dos casos ou ainda com uma eventual escolha de doentes por parte dos hospitais.

A consideração da primeira alternativa tem essencialmente em atenção aspectos da rede hospitalar portuguesa, na qual a lógica prevalecente poderá eventualmente residir numa concentração de recursos para se tratarem os casos mais complexos ou mais graves.

No capítulo da apresentação de resultados foi referido que o nível de produção tem efeitos sobre a complexidade dos casos tratados, aumentando com incrementos da produção, sobre a gravidade dos casos tratados, diminuindo com incrementos da produção e sobre a diferença entre a complexidade e a gravidade dos casos tratados, igualmente aumentando com incrementos da produção.

Deve igualmente ser evidenciado que esta análise global e genérica pode ser, pelo menos parcialmente infirmada quando se realizar uma avaliação mais desagregada, essencialmente por doença principal, pelo que no futuro se sugere que este tipo de estudos sejam realizados em Portugal, para que se possa caracterizar o comportamento dos hospitais, conhecer as suas razões justificativas e introduzir mecanismos para se corrigir os problemas detectados.

Enquanto, tal não for realizado, deve ter-se presente que este estudo indicia um comportamento distinto dos hospitais, no qual as organizações de saúde, com maior dimensão são as que simultaneamente tratam doentes mais complexos e menos graves.

Estas conclusões não eram até ao presente momento reconhecidas em Portugal, visto que se defendia que os hospitais Universitários, igualmente aqueles que na generalidade apresentam maior capacidade de oferta e maior dimensão produtiva, por estarem no fim da linha da referência hospitalar, seriam os que tratavam doentes mais graves (Serrão e outros, 1998).

Este estudo mostra que a gravidade dos doentes tratados nos hospitais de maior dimensão, tem sido confundida por outros aspectos, designadamente a raridade dos casos tratados e ainda a sua complexidade.

No entanto, o estudo aponta ainda para o facto de os hospitais com doentes menos graves, menos complexos e menos graves e complexos apresentarem melhores resultados brutos, quando os mesmos estão a ser medidos pela mortalidade e pela demora média.

Este estudo evidencia ainda que quanto maior é a diferença entre a complexidade e a gravidade dos casos tratados melhores são os resultados dos hospitais, expressos por taxas de mortalidade e demoras média mais baixas.

Esta situação é a esperada face a diversos estudos internacionais, onde são evidenciadas as vantagens de admitir doentes com menor gravidade e complexidade, não só ao nível dos resultados, mas igualmente sobre o desempenho financeiro dos hospitais (Hornbrook e Goldfarb, 1983; Gonnella, Hornbrook e Louis, 1984; Schumacher et al, 1987; Thomas e Ascraft, 1989 e Barnum, Kutzin e Saxenian, 1995).

Neste sentido procedeu-se a outra análise em que se associava o perfil das admissões com os resultados ajustados pelo risco.

Como grande conclusão deve referir-se que não existe nenhuma associação entre a gravidade, a complexidade, a gravidade e complexidade e a diferença entre estes dois atributos e o desempenho dos hospitais, tanto para a efectividade, como para a eficiência e naturalmente para os resultados globais.

Esta situação é completamente diferente da retratada quando se consideram somente os resultados brutos, visto que genericamente os hospitais apresentavam melhor desempenho quando tratavam doentes menos graves e complexos ou maximizavam a diferença entre estes dois atributos.

Estas conclusões que podem ser consideradas esperadas, atendendo a que os hospitais estavam classificados de uma forma distinta quando se consideravam os resultados brutos e os resultados ajustados pelo risco.

Nesta conformidade, pode referir-se que a introdução do ajustamento pelo risco, elimina pelo menos um dos problemas decorrentes de diferentes políticas de admissão por parte dos hospitais – a apresentação de melhores indicadores e de resultados da sua actividade.

Para a avaliação do desempenho dos hospitais privilegiou-se a utilização de indicadores de resultados em função de três dimensões: a efectividade, a eficiência e o desempenho global.

Em relação à efectividade dos cuidados prestados foram retiradas as seguintes conclusões:

- Existem indicações completamente distintas quando se analisa o desempenho dos hospitais em função dos seus resultados brutos e ajustados pelo risco. Estas conclusões são válidas para qualquer das perspectivas de análise, ou seja, todas as admissões, casos cirúrgicos e médicos e por doença;
- Observou-se ainda um comportamento bastante heterogéneo dos hospitais quando a avaliação do desempenho é feita de uma forma mais desagregada. Ou seja, os hospitais em termos genéricos apresentam comportamentos distintos entre os casos cirúrgicos e médicos e, de uma forma mais nítida, por doença.

Estes aspectos merecem algumas reflexões suplementares:

- Em primeiro lugar, deve referir-se que os argumentos apresentados pelos hospitais que os seus resultados brutos não devem ser comparados, essencialmente devido às características dos seus doentes, onde avulta a gravidade, parecem ter cabimento;
- Em segundo lugar, deve evidenciar-se o desempenho distinto dos hospitais, tanto entre os casos cirúrgicos e médicos, como por doença. A este propósito foi inclusivamente referido anteriormente que existem hospitais cujas diferenças no desempenho por doença são por vezes superiores a diferenças encontradas entre diferentes hospitais.

Por outro lado, um ensaio explicativo sobre o desempenho dos hospitais no que se refere à efectividade, para o qual foram utilizadas as variáveis independentes, Dimensão, Tipo de hospital e Localização do hospital, permite as seguintes conclusões:

- Existe um efeito dimensão da produção para os resultados ajustados pelo risco, tanto para o total das admissões, como para os casos cirúrgicos e médicos;
- O ajustamento que melhor traduz esta relação assume uma forma quadrática, numa forma de U invertido, pelo que a melhor efectividade é encontrada nos hospitais com maiores e menores volumes de produção, embora com maior intensidade para a maior produção. Esta afirmação é igualmente válida para o total das admissões e para os casos cirúrgicos e médicos;
- A localização dos hospitais, medida pela Região de Saúde, tem igualmente influência para o total das admissões e para os casos médicos;
- Para os casos cirúrgicos esta dimensão não está presente, sendo no entanto relevante o Tipo de hospital.

Esta análise macro, embora identifique a relevância da produção para a obtenção de níveis de efectividade não é totalmente conclusiva para efeitos de

planeamento e da administração das organizações de saúde, visto que os melhores valores podem ser obtidos com altos e baixos níveis de efectividade.

Apesar destas conclusões, parece igualmente evidente que os hospitais com maior dimensão produtiva são os mais efectivos, pelo que se torna importante a realização de estudos mais desagregados, essencialmente ao nível da doença, para que desta forma se possa perspectivar o desenho futuro dos serviços e dos hospitais em Portugal.

Atendendo a que os resultados dos hospitais para casos cirúrgicos e médicos foram distintos, procedeu-se a uma análise tentando investigar se este fenómeno era estrutural. A principal conclusão aponta para que, embora os hospitais sejam classificados de forma diferente segundo o seu desempenho para os casos cirúrgicos e médicos, tudo parece indicar que este fenómeno não é estrutural, visto que, na generalidade, os hospitais que são mais efectivos nos casos cirúrgicos, também o são nos casos médicos.

No entanto, demonstrou-se igualmente neste estudo que parece existir um efeito dimensão dos hospitais para a justificação do diferente desempenho entre casos médicos e cirúrgicos, com especial incidência nas organizações de saúde com maior dimensão, tanto em termos do total de episódios de internamento, como na percentagem de casos cirúrgicos tratados.

Finalmente, em relação às grandes discrepâncias encontradas no desempenho dos hospitais, pode afirmar-se que praticamente não existe nenhum padrão para a existência de discrepâncias no desempenho interno dos hospitais, no que respeita à efectividade.

Em relação à eficiência dos cuidados prestados foram retiradas as seguintes conclusões:

- Existem indicações completamente distintas quando se analisa o desempenho dos hospitais em função dos seus resultados brutos e ajustados pelo risco. Estas conclusões são válidas para qualquer das perspectivas de análise, ou seja, todas as admissões, casos cirúrgicos e médicos e por doença;
- Observou-se ainda um comportamento bastante heterogéneo dos hospitais quando a avaliação do desempenho é feita de uma forma mais desagregada. Ou seja, os hospitais em termos genéricos apresentam comportamentos distintos entre os casos cirúrgicos e médicos e, de uma forma mais nítida, por doença.

Estas conclusões são praticamente idênticas às encontradas quando se analisa a efectividade dos hospitais, reforçando assim os argumentos apresentados pelos hospitais, os quais parecem ter cabimento, visto que os seus resultados

brutos não devem ser comparados, essencialmente devido às características dos seus doentes, onde avulta a gravidade.

À semelhança do realizado para a efectividade, ensaiou-se um modelo explicativo sobre o desempenho dos hospitais no que se refere à eficiência, para o que foram utilizadas as variáveis independentes, Dimensão, Tipo de hospital e Localização do hospital, permitindo as seguintes conclusões:

- Existe um efeito dimensão da produção para os resultados ajustados pelo risco somente para os casos cirúrgicos;
- Para estas situações o ajustamento que melhor traduz esta relação assume uma forma quadrática, numa forma de U, pelo que a melhor eficiência é encontrada nos hospitais com dimensões médias;
- Para o total das admissões e para os casos médicos, as variáveis incluídas no modelo, não justificam o desempenho dos hospitais.

Esta análise macro, embora não sendo totalmente conclusiva para justificar os diversos níveis de eficiência, permite desde já identificar os seguintes aspectos:

- Comportamento distinto entre o total das admissões e os casos médicos e os casos cirúrgicos;
- Pouca relevância da dimensão e das características da oferta no desempenho dos hospitais;
- Comportamento distinto entre a efectividade e a eficiência dos hospitais.

Mais uma vez foram encontradas grandes discrepâncias na eficiência intra-hospitalar, pelo que se reitera a necessidade de se melhorar a gestão e a regulação dos hospitais.

No que se refere ao desempenho dos hospitais para a efectividade e para a eficiência, tanto na apresentação dos resultados, como na discussão sobre estas dimensões, foi referida a existência de indicações contrárias e potencialmente contraditórias.

Na realidade, a ordenação dos hospitais medida pelo *K de Cohen* permite concluir pela existência de um comportamento completamente diferente dos hospitais na efectividade e na eficiência.

Por outro lado, os valores não significativos do coeficiente de correlação de Spearman (ordenação de hospitais) e de Pearson ("z scores") entre a efectividade e a eficiência dos hospitais, tanto para o total das admissões, como para casos cirúrgicos e médicos, parecem igualmente traduzir a inexistência de uma associação entre estas duas dimensões.

No entanto, atendendo por um lado, a que todas as associações são positivas e a que, por outro lado, nunca se observam correlações significativas e com sinal negativo, pode concluir-se que embora o desempenho dos hospitais seja diferente para a efectividade e para a eficiência, não existem contradições entre estas duas dimensões. Ou seja, para se otimizar a efectividade não é necessário sacrificar a eficiência, sendo o inverso igualmente verdadeiro.

Esta análise é ainda parcialmente verificada quando se retiram os hospitais com maiores diferenças entre a efectividade e a eficiência.

Assim, embora para todos os hospitais não exista qualquer associação entre a efectividade e a eficiência, tal não corresponde a um fenómeno estrutural em Portugal, visto que com a exclusão dos 24 hospitais com grandes discrepâncias entre estas duas dimensões ou somente dos 12 hospitais com maiores níveis de efectividade em relação à eficiência, a correlação é directa e significativa. Ou seja quanto maior a efectividade maior é a eficiência e vice-versa.

Para o modelo global de avaliação do desempenho hospitalar (modelo misto), no qual são incluídas a efectividade e a eficiência dos cuidados prestados, foram discutidas previamente duas questões sobre a sua operacionalização: utilização das ordenações dos hospitais ou dos valores directos do seu desempenho e a ponderação que deve ser atribuída a cada uma das dimensões.

Apesar das indicações serem distintas, neste estudo optou-se pela utilização directa dos valores da efectividade e da eficiência e com igual ponderação destas duas dimensões, tendo-se chegado às seguintes conclusões:

- Nos 10 melhores hospitais encontram-se 5 tanto do Tipo I como do III, sendo 9 comuns à da classificação da efectividade e 5 à da eficiência;
- Para os 10 piores hospitais encontra-se uma maior concentração pela tipologia, visto que 8 são do Tipo III, sendo os restantes 2 dos Tipos I e V;
- No entanto, na globalidade, os hospitais são classificados de forma diferente, tanto da efectividade, como da eficiência, sendo a concordância na respectiva classificação baixa;
- À semelhança do que foi encontrado para as análises da efectividade e da eficiência, os hospitais apresentam comportamento distinto para os casos médicos e por doença. Para esta última perspectiva é igualmente afirmado que, por vezes, existem maiores diferenças intra-hospitalares do que entre hospitais.

Para o desempenho global foi igualmente ensaiado um modelo explicativo sobre o desempenho dos hospitais, para o qual foram utilizadas as variáveis independentes, Dimensão, Tipo de hospital e Localização do hospital, permitindo as seguintes conclusões:

-
- Para o total das admissões a única variável explicativa é a localização dos hospitais (com especial relevo para as Regiões Centro e Norte);
 - Para os casos cirúrgicos o modelo não é explicativo, tanto globalmente, como por variáveis;
 - Para os casos médicos, todas as variáveis são significativas. Para a produção os hospitais com melhor desempenho são encontrados entre aqueles com maior e menor dimensão (embora com maior relevo no primeiro conjunto de hospitais);
 - Para qualquer destes modelos o poder explicativo global do modelo (R^2 ajustado) é sempre menor do que o encontrado para os resultados brutos.

Estes aspectos comprovam genericamente que a utilização dos resultados brutos para se avaliar o desempenho dos hospitais conduz a conclusões erróneas, pelo que a introdução do ajustamento pelo risco constitui um instrumento necessário para se concretizar esta finalidade.

Por outro lado, para além do diferente comportamento dos hospitais entre os casos médicos e cirúrgicos, constata-se que não existem tipos de hospital, nem dimensões óptimas para se maximizarem os resultados.

Assim, pese embora o facto de as Regiões Centro e Norte apresentarem resultados melhores, deve concluir-se que em Portugal o melhor ou pior desempenho, parece estar mais associado com a gestão dos hospitais do que com características intrínsecas das organizações de saúde.

Contudo, face ao exposto quando se analisa o desempenho por doença para qualquer das dimensões consideradas, no qual avulta a existência de grandes disparidades intra-hospitalares e de diversos hospitais com melhor ou pior desempenho para cada uma das doenças, deve ainda reforçar-se que este esforço da gestão é mais patente ao nível do serviço do que do próprio hospital.

Ou seja, estes resultados da investigação parecem exprimir que em cada hospital existem vários “hospitais” cada um com a sua dinâmica própria, pelo que se pode ainda concluir que ainda existe uma insuficiência ao nível macro ou estratégico da gestão dos hospitais.

Se tal se deve somente a deficiências das administrações hospitalares, ou deve ser atribuído ao actual enquadramento normativo e modelo de financiamento ou a um “mix” destas situações são questões que não são alvo deste estudo. No entanto, atendendo à importância do fenómeno, bem como ao processo de mudança organizacional em curso, sugere-se que no futuro sejam realizados estudos específicos sobre estes assuntos.

Neste estudo também foram evidenciadas algumas implicações dos resultados na reorganização do sector da saúde. Esta atitude parte de princípio que embora a avaliação da produção e do desempenho hospitalar, tenha um valor intrínseco, não constitui somente um fim, mas igualmente um ponto de partida para se identificarem problemas, para se definirem prioridades e para se introduzirem melhorias no sistema.

Nesta conformidade, foram discutidos aspectos relacionados com o financiamento e com a regulação e gestão dos hospitais e ainda com a liberdade de escolha por parte dos consumidores.

No que se refere ao financiamento foram feitas as seguintes recomendações:

- A definição de um mecanismo de preços com base em custos por doença, sendo para tal necessário que se desenvolva um sistema de custeio por doente;
- A introdução de níveis de gravidade para diferenciar preços dentro do mesmo produto. Para tal podem ser utilizadas as indicações do Disease Staging, ou de uma das versões dos DRGs que incluem a gravidade (APRDRGs ou IRDRGs) ou ainda de um sistema que utilize a informação clínica;
- Definição da demora média central por DRG não em função do que foi feito, mas da demora média esperada. Mais uma vez a utilização dos sistemas administrativos ou clínicos podem introduzir informações válidas para o efeito.

Para além destas questões mais técnicas do financiamento, foi igualmente sugerido que se introduzissem outras questões de política de saúde, nomeadamente a definição de um sistema de incentivos.

A este respeito sugeriu-se que, na eventualidade de existir um sistema de pagamento por caso, este seja replicado pelos profissionais de saúde e que se criem preços diferenciais, não em função do tipo de hospital, mas de níveis de desempenho.

Para a função regulação foram analisadas duas vertentes: o perfil das admissões e os diferentes níveis de desempenho entre os hospitais.

No que se refere ao perfil das admissões, atendendo a que a mobilidade dos doentes não segue um modelo linear, visto que os doentes que são tratados fora do local da residência são simultaneamente mais complexos e menos graves e a que se encontraram grandes diferenças no perfil das admissões nos hospitais portugueses, no qual os hospitais de maior dimensão foram os que maximizaram a diferença entre a complexidade e a gravidade dos casos tratados, com maior expressão no primeiro atributo, recomenda-se igualmente o seguinte:

-
- A realização de estudos sobre a conformidade destes comportamentos com os princípios e finalidades do Serviço Nacional de Saúde Português;
 - A definição de orientações sobre a referência de doentes, tanto a montante, como a jusante;
 - A definição de mecanismos e de indicadores para monitorizar a mobilidade dos doentes;
 - A definição de uma política de incentivos (incluído penalizações) para premiar ou punir os hospitais e/ou prestadores que apresentem um comportamento não esperado.

Para o diferente desempenho recomenda-se o seguinte:

- A realização de estudos mais detalhados para se investigarem as causas para as diferenças apontadas no desempenho dos hospitais do mesmo Tipo e/ou da mesma Região de Saúde;
- A definição de políticas e de medidas que contribuam para a redução nas disparidades no desempenho hospitalar;
- A criação de um sistema de informação para monitorizar o desempenho dos hospitais.

No que respeita à gestão dos hospitais, atendendo a que as disparidades internas, tanto no perfil das admissões, como no desempenho em termos de efectividade, de eficiência e global, são por vezes mais marcantes que as encontradas foi referida a insuficiência na prática da gestão hospitalar, passando, na generalidade, a sua melhoria com medidas do mesmo tipo, embora com intensidade diferente de aplicação, das que foram referidas para a função regulação.

Finalmente, defendeu-se que os resultados do desempenho dos hospitais devem ser sempre publicitados e que esta actividade ganha ainda maior relevo na eventualidade de se pretender implementar a liberdade de escolha por parte dos consumidores.

Em relação à liberdade de escolha foram muito genericamente abordados dois modelos: escolha do hospital ou escolha por tipo de cuidados de saúde, referindo-se que esta última alternativa pode parecer mais atractiva para os consumidores.

No entanto, referiu-se ainda que existem potenciais conflitos entre a liberdade de escolha por tipo de cuidados de saúde e a efectividade e a eficiência das organizações de saúde, pelo que se devem ter cuidados suplementares na sua implementação, essencialmente ao nível do sistema de informação, nomeadamente através de uma maior informatização e de maior flexibilidade nos seus fluxos de disseminação.

Como síntese final evidencia-se que, pese embora o facto de a avaliação do desempenho dos hospitais e eventualmente a sua publicitação devam passar a constituir um objectivo prioritário do Sistema Nacional de Saúde, algumas cautelas devem ser tomadas.

Em primeiro lugar, associadas com a escolha dos indicadores, visto que para além da justificação da respectiva pertinência, a sua inclusão ou omissão e ainda distinta valorização pode conduzir a avaliações erróneas sobre o desempenho dos hospitais.

Em segundo lugar, que os critérios de avaliação do desempenho não incluam apenas indicadores associados com a oferta, visto que em última instância a finalidade de todos os hospitais, independentemente do seu nível de desenvolvimento tecnológico é tratar doentes.

Por último, as características dos doentes devem ser consideradas, nomeadamente a gravidade e o consequente ajustamento pelo risco, o qual para além de minimizar os problemas decorrentes do diferente “mix” dos doentes tratados, permite ainda a comparação entre valores observados e valores esperados, o que se reitera deve ser uma metodologia a implementar em Portugal.

BIBLIOGRAFIA

Bibliografia

- Australian Council on Health Care Standards, 2002. Determining the Potential to Improve Quality of Care – ACHS Indicator Results for Australia and New Zealand 1998-2002. The ACHS, Health Services Research Group, University of Newcastle, Ultimo, New South Wales, 2002.
- Agência de Contratualização dos Serviços de Saúde, 2002. Contratualização com os Hospitais para 2003. Ministério da Saúde, Região de Saúde de Lisboa e do Vale do Tejo, Lisboa, 2002.
- Agência de Contratualização dos Serviços de Saúde, 2002. Projecto de Indicadores. Ministério da Saúde, Região de Saúde de Lisboa e do Vale do Tejo, Lisboa, 2002.
- Ahicart C, 1988. Técnicas de Medicion del Case-Mix Hospitalario. Hospital 2000, Suplementos 1, 2, 3 e 4.
- Almeida RT e Carlsson P, 1996. Severity of a Case for Outcome Assessment in Health Care – Definitions and Classifications of Instruments. Health Policy, 37 (1), 35-52.
- Altman D, Cutler D e Zeckhauser R, 2003. Enrollee Mix, Treatment Intensity, and Cost in Competing Indemnity and HMO Plans. Journal of Health Economics, 22 (1), 23-45.
- Alves AD, 1994. Avaliação da Performance dos Hospitais Portugueses. Tese de Mestrado, Instituto Superior de Estudos Empresariais da Universidade do Porto, Porto, 1994.
- Amaratunga D, Haigh R, Sarshar M e Baldry D, 2002. Application of the Balanced Scorecard Concept to Develop a Conceptual Framework to Measure Facilities Management Performance within NHS Facilities. International Journal of Health Care Quality Assurance, 15 (4), 141-151.
- Ament RP, Dreachslin JL, Kobrinsky EJ e Wood WR, 1982. Three Case-Type Classifications: Suitability for the Use in Reimbursing Hospitals. Medical Care, 20 (5), 460-467.
- Anthony RN e Herzlinger RE, 1975. Management Control in Nonprofit Organizations. Richard D. Irwin, Inc, Illinois, 1975.
- Applegate WB, Blass JP e Williams TF, 1990. Instruments for the Functional Assessment of the Older Patients. New England Journal of Medicine, 322 (17), 1207-1214.
- Arndt M, Bradbury RC e Golec JH, 1995. Surgeon Volume and Hospital Resource Utilization. Inquiry, 32 (4), 407-417.
- Aronow DB, 1988. Severity of Illness Measurement: Applications in Quality Assurance and Utilization Review. Medical Care Review, 45 (2), 339-366.
- Arozullah AM, Henderson WG, Khuri SF e Daley J, 2003. Postoperative Mortality and Pulmonary Complication Rankings: How Well Do They Correlate at the Hospital Level? Medical Care, 41 (8), 979-991.
- Ash AA e Schwartz M, 1997. Evaluating the Performance of Risk-Adjustment Methods: Dichotomous Variables. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 427-469.
- Averill RF, Muldoon JH, Vertrees JC, Goldfield NI, Mullin RL, Fineran EC, Zhang MZ, Steinbeck B e Grant T, 1998. The Evolution of Casemix Measurement Using Diagnosis Related Groups (DRGs). 3M Health Information Systems, 5-98.
- Barer ML, 1982. Case Mix Adjustment in Hospital Cost Analysis: Information Theory Revisited. Journal of Health Economics, 1 (1), 53-80.
- Barnum H, Kutzin S e Saxenian H, 1995. Incentives and Provider Payment Methods. International Journal of Health Planning & Management, 10 (1), 23-45.

-
- Barros PP, 2001a. Um Exemplo Simples da Metodologia de Categorização de Hospitais segundo “Grade of Membership”. In Instituto de Gestão Informática e Financeira, “Orçamento do SNS para 2002 – Financiamento dos Hospitais”, Lisboa, 2001.
 - Barros PP, 2001b. Utilização dos Graus de Pertença Para Cálculo de Orçamentos dos Hospitais, In Instituto de Gestão Informática e Financeira, “Orçamento do SNS para 2002 – Financiamento dos Hospitais”, Lisboa, 2001.
 - Barros PP, 2001c. Exploração Preliminar da Aplicação da Metodologia de Graus de Pertença. In Instituto de Gestão Informática e Financeira, “Orçamento do SNS para 2002 – Financiamento dos Hospitais”, Lisboa, 2001.
 - Barros PP, 2001c. Exploração Preliminar da Aplicação da Metodologia de Graus de Pertença. In Instituto de Gestão Informática e Financeira, Orçamento do SNS para 2002 – Financiamento dos Hospitais, Lisboa, 2001.
 - Becker ER e Steinwald B, 1991. Determinants of Hospital Casemix Complexity. *Health Services Research*, 16 (4), 439-458.
 - Becker RB, Zimmerman JE, Knaus WA, Wagner DP, Seneff MG, Draper EA, Higgins TL, Estafanous FG e Loop FD, 1995. The Use of APACHE III to Evaluate ICU Length of Stay, Resource Use, and Mortality after Coronary Artery By-pass Surgery. *Journal of Cardiovascular Surgery*, 36 (1), 1-11.
 - Bennet CL, Greenfield H, Aronow P, Ganz P, Vogelzang NJ e Elashoff RM, 1991. Patterns of Care Related to Age of Men with Prostate Cancer. *Cancer*, 67 (10), 2633-2641.
 - Bentes M, Gonçalves M, Tranquada S e Urbano J, 1996. A Utilização de GDH's como Instrumento de Financiamento Hospitalar. *Gestão Hospitalar*, 33 (9), 33-40.
 - Bentley JD e Butler PW, 1981. The DRG Case Mix of a Sample of Teaching Hospitals: A Technical Report. Association of Medical Colleges, Washington, 1981.
 - Bentley JD, Butler PW, 1980. Case-Mix Reimbursement: Measures, Applications, Experiments. *Hospital Financial Management*, 34 (3), 24-35.
 - Berki SE, 1972. *Hospital Economics*. Lexington Books, Lexington, Massachussets, 1972.
 - Berki SE, Ashcraft ML e Newbrander WC, 1984. Length-of-Stay Variations within ICDA-8 Diagnosis Related Groups. *Medical Care*, 22 (2), 126-142.
 - Bindman AB e Grumbach K, 1992. America's Safety Net. The Wrong Place at the Wrong Time. *JAMA*, 268 (17), 2426-2427.
 - Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, Welch HG e Wennberg DE, 2002. Hospital Volume and Surgical Mortality in the United States. *New England Journal of Medicine*, 346 (15), 1128-1137.
 - Bleichrodt H, Crainich D e Eeckhoudt L, 2003. The Effect of Comorbidities on Treatment Decisions. *Journal of Health Economics*, 22 (5), 805-820.
 - Blumberg MS, 1986. Risk Adjusting Health Care Outcomes: A Methodologic Review. *Medical Care Review*, 43 (2), 351-393.
 - Blumberg MS, 1987. Comments on HCFA Hospital Death Rate Statistical Outliers. *Health Services Research*, 21 (6), 715-739.
 - Bradbury RC, Golec JH e Steen PM, 1994. Relating Hospital Health Outcomes and Resource Expenditures. *Inquiry*, 31 (1), 56-65.
 - Brailer DJ, Kroch E, Pauly MV e Huang J, 1996. Comorbidity-Adjusted Complication Risk – A New Outcome Quality Measure. *Medical Care*, 34 (5), 490-505.
 - Bravo G, Dubois MF, Wals P, Herbert R e Messier L, 2002. Relationship between Regulatory Status, Quality of Care, and Three-Year Mortality in Canadian Residential Care Facilities: A Longitudinal Study. *Health Services Research*, 37 (5), 1181-1196.
 - Brewster AC, Jacobs CM e Bradbury RC, 1984. Classifying Severity of Illness by Using Clinical Findings. *Health Care Financing Administration, Annual Supplement*, 107-108.

-
- Brewster AC, Karlin BG, Hyde LA, Jacobs CM, Bradbury RC e Chae YM, 1985. MEDISGROUPS: A Clinically Based Approach to Classifying Patients at Hospital Admission. *Inquiry*, 22 (4), 377-387.
 - Brook RH, Iezzoni LI, Jencks SF, Knaus WA, Krakauer H, Lohr KN e Moskowitz MA, 1987. Symposium: Case-Mix Measurement and Assessing Quality of Hospital Care. *Health Care Financing Review*, December (Special Number), 39-48.
 - Burns LR, Wholey DR e Abeln MO, 1993. Hospital Utilization and Mortality Levels for Patients in the Arizona Health Care Cost Containment System. *Inquiry*, 30 (2), 142-156.
 - Burstin HR, Lipsitz SR e Brennan TA, 1992. Socioeconomic Status and Risk for Substandard Medical Care. *JAMA*, 268 (17), 2383-2387.
 - Butler JR, 1995. *Hospital Cost Analysis*. Kluwer Academic Publishers, Netherlands, 1995.
 - Cabral J e Barriga N, 1999. Economias de Escala, Eficiência e Custos nos Hospitais Distritais. Evidências Empíricas. Associação Portuguesa de Economia da Saúde, Documento de Trabalho 2/99, Lisboa, 1999.
 - Calore KA e Iezzoni LI, 1987. Disease Staging and PMCs – Can They Improve DRGs? *Medical Care*, 25 (8), 724-735.
 - Cao Z e McGuire TG, 2003. Service-Level Selection by HMOs in Medicare. *Journal of Health Economics*, 22 (6), 915-931.
 - Carneiro AV, 1994. Factores de Prognóstico em Medicina Intensiva – Aplicação de Sistemas Gerais de Prognóstico em Três Unidades de Cuidados Intensivos de Um Hospital Universitário. Tese de Doutorado, Faculdade de Medicina de Lisboa, Universidade de Lisboa, Lisboa, 1994.
 - Carreira CM, 1999. Economias de Escala e de Gama nos Hospitais Públicos Portugueses: Uma Aplicação da Função de Custo Variável Translog. Associação Portuguesa de Economia da Saúde, Documento de Trabalho 3/99, Lisboa, 1999.
 - Casas M, 1991. Clasificación de Pacientes y Producción Hospitalaria: Los Grupos Relacionados con el Diagnóstico (GRD). In Casas M (editor), “Los Grupos Relacionados con el Diagnóstico (GRD): Experiencia y Perspectiva de Utilización”, Masson, Barcelona, 1991, 23-43.
 - Case RB, Moss AJ, Case N, McDermott M e Eberly S, 1992. Living Alone After Myocardial Infarction: Impact on Prognosis. *JAMA*, 267 (4), 515-519.
 - Chang LC, Lin SW e Northcott DN, 2002. The NHS Performance Assessment Framework – A “Balanced Scorecard” Approach? *Journal of Management in Medicine*, 16 (5), 345-358.
 - Charbonneau C, Ostrowsky C, Phoener ET, Lindsay P, Panniers TL, Houghton P, Albright J, 1988. Validity and Reliability Issues in Alternative Patient Classification Systems. *Medical Care*, 26 (8), 800-813.
 - Charlson ME, Pompei P, Ales KL e MacKenzie CR, 1987. A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation. *Journal of Chronic Diseases*, 40 (5), 373-383.
 - Chassin MR, Park RE, Lohr KL, Keesey J e Brook, RH, 1989. Differences among Hospitals in Medicare Patient Mortality. *Health Services Research*, 24 (1), 1-31.
 - Chen E e Naylor D, 1994. Variation in Hospital Length of Stay for Acute Myocardial Infarction in Ontario, Canada. *Medical Care*, 32 (5), 420-435.
 - Clark DE e Ryan LM, 2002. Concurrent Prediction of Hospital Mortality and Length of Stay from Risk Factors on Admission. *Health Services Research*, 37 (3), 631-645.
 - Cleary PD, Greenfield S, Mulley AG, Pauker SG, Schroeder SA, Wexler L e McNeil BJ, 1991. Variations in Length of Stay and Outcomes for Six Medical and Surgical Conditions in Massachusetts and California. *JAMA*, 266 (1), 73-79.
 - Cleves MA, Sanchez N e Draheim M, 1997. Evaluation of Two Competing Methods for Calculating Charlson's Comorbidity Index when Analyzing Short-Term Mortality Using Administrative Data. *Journal of Clinical Epidemiology*, 50 (8), 903-908.

-
- Coffey RM e Goldfarb, MG, 1986. DRGs and Disease Staging for Reimbursing Medicare Patients. *Medical Care*, 24 (9), 814-829.
 - Conklin JE, Lieberman JV, Barnes CA e Louis DZ, 1984. Disease Staging: Implications for Hospital Reimbursement and Management. *Health Care Financing Administration, Annual Supplement*, 13-21.
 - Costa C, Delgado M e Carvalho R, 1989. Relação entre as Características da Procura e a Utilização de Recursos num Hospital. *Revista Portuguesa de Saúde Pública*, 7 (4), 27-40.
 - Costa C e Reis V, 1993. O Sucesso nas Organizações de Saúde. *Revista Portuguesa de Saúde Pública*, 11 (3), 59-68.
 - Costa C e Nogueira P, 1994. Produção Hospitalar e Fiabilidade. *Revista Portuguesa de Saúde Pública*, 12 (2), 31-40.
 - Costa C e Sena C, 2002. Análise de Alguns Indicadores de Funcionamento dos Hospitais no Período 1990-1999. Documento de Trabalho 1/2002, Curso de Especialização em Administração Hospitalar, Escola Nacional de Saúde Pública, Lisboa, 2002.
 - Costa C, 1990. Financiamento de Serviços de Saúde – A Definição de Preços. *Revista Portuguesa de Saúde Pública*, 8 (2), 65-72.
 - Costa C, 1991. A Severidade da Doença – Identificação e Caracterização de Alguns Sistemas de Classificação. *Revista Portuguesa de Saúde Pública*, 9 (1), 37-43.
 - Costa C, 1994. Os DRGs (Diagnosis Related Groups) e a Gestão do Hospital. *Revista Portuguesa de Gestão*, III/IV, 47-65.
 - Culler S e Ehrenfried D, 1986. On the Feasibility and Usefulness of Physician DRGs. *Inquiry*, 23 (1), 40-55.
 - Daley, 1997. Validity of Risk-Adjustment Methods. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 330-363.
 - Daley J, Jencks S, Draper D, Lenhart G, Thomas N e Walker J, 1988. Predicting Hospital-Associated Mortality for Medicare Patients: A Method for Patients with Stroke, Pneumonia, Acute Myocardial Infarction and Congestive Heart Failure. *JAMA*, 260 (24), 3617-3624.
 - Davis RB, Iezzoni LI, Philips RS, Reiley P, Coffman GA e Safran C, 1995. Predicting In-Hospital Mortality – The Importance of Functional Status Information. *Medical Care*, 33 (9), 906-921.
 - Degeling P, Sorensen R, Maxwell S, Aisbet C, Zhang K e Coyle B, 2000. The Organization of Hospital Care and its Effects. Centre for Clinical Governance Research, University of New South Wales, Sydney, 2000.
 - DesHarnais SI, Forthman MT, Lowry JM e Wooster LD, 1997. Risk-Adjusted Quality Outcome Measures: Indexes for Benchmarking Rates of Mortality, Complications and Readmissions. *Quality Management in Health Care*, 5 (2), 80-87.
 - DesHarnais SI, Forthman MT, Lowry JM e Wooster LD, 2000. Risk-Adjusted Clinical Quality Indicators: Indexes for Measuring and Monitoring Rates of Mortality, Complications, and Readmissions. *Quality Management in Health Care*, 9 (1), 14-22.
 - Deyo RA, Cherkin DC e Ciol MA, 1992. Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Databases. *Journal of Clinical Epidemiology*, 45 (6), 613-619.
 - Direcção Geral de Saúde, 2002. Desempenho Comparado das Unidades de Saúde do SNS – Indicadores Mensais, 2002. DGS, Lisboa, 2002.
 - Dismuke CE e Sena V, 1998. Hospital Productivity and Efficiency Measurement in the Presence of Undesirable Output. Associação Portuguesa de Economia da Saúde, Documento de Trabalho 2/98, Lisboa, 1998.
 - Dismuke C, 1996. O Impacto do Sistema de GDH na Eficiência dos Hospitais em Portugal. Escola de Economia e Gestão da Universidade do Minho, Braga, 1996.
 - Donabedian A, 1980. Methods for Deriving Criteria for Assessing the Quality of Medical Care. *Medical Care Review*, 37 (7), 653-698.
 - Donabedian A, 1985. The Epidemiology of Quality. *Inquiry*, 22 (3), 282-292.

-
- Donabedian A, 1986. Criteria and Standards for Quality Assessment and Monitoring. *Quality Review Bulletin*, 12 (3), 99-108.
 - Dranove D, 1998. Economies of Scale in Non-Revenue Producing Cost Centers: Implications for Hospital Mergers. *Journal of Health Economics*, 17 (1), 69-83.
 - Dubois RW, Brook RH e Rogers WH, 1987. Adjusted Hospital Death Rates: A Potential Screen for Quality of Medical Care. *American Journal of Public Health*, 77 (9), 1162-1166.
 - Dubois RW, Rogers WH, Moxley JH, Draper D e Brook RH, 1987. Hospital Inpatient Mortality – Is it a Predictor of Quality? *The New England Journal of Medicine*, 317 (26), 1674-1680.
 - Elixhauser A, Steiner C, Harris DR e Coffey RM, 1998. Comorbidity Measures for Use with Administrative Data. *Medical Care*, 36 (1), 8-27.
 - Ellis RP, 1998. Creaming, Skimping and Dumping: Provider Competition on the Intensive and Extensive Margins. *Journal of Health Economics*, 17 (5), 537-555.
 - Emerson JD e Strenio J, 1983. Caixas de Bigodes e Comparação de Coleções. In Hoaglin DC, Mosteller F e Tukey JW (editores), “Análise Exploratória de Dados – Técnicas Robustas”, Edições Salamandra, Lisboa, 59-96.
 - Epstein AM, Boge J, Dreyer P e Thorpe KE, 1991. Trends in Length of Stay and Rates of Readmission in Massachusetts: Implications for Monitoring Quality of Care. *Inquiry*, 28 (1), 19-28.
 - Evans JH, Hwang Y e Nagarajan N, 1995. Physicians’ Response to Length-of-Stay Profiling. *Medical Care*, 33 (11), 1106-1119.
 - Evans RG e Walker HD, 1972. Information Theory and the Analysis of Hospital Cost Structure. *Canadian Journal of Economics*, 5, 398-418.
 - Evans, RG, 1981. Incomplete Vertical Integration: The Distinctive Structure of Health Care Industry. In Van der Gaag J e Perlman M (editors), “Health, Economics and Health Economics”, North Holland, Amsterdam, 1981, 329-354.
 - Evans RG, 1984. *Strained Mercy: The Economics of Canadian Health Care*. Butterworths, Toronto, 1984.
 - Farley DE e Hogan C, 1990. Case-Mix Specialization in the Market for Hospital Services. *Health Services Research*, 25 (5), 757-783.
 - Farley DE e Ozminkowsky RJ, 1992. Volume-Outcome Relationships and Inhospital Mortality: The Effect of Changes in Volume Over Time. *Medical Care*, 30 (1), 77-94.
 - Fetter RB, Shin Y, Freeman JL, Averill RF e Thompson JD, 1980. Case Mix Definition by Diagnosis-Related Groups. *Medical Care*, 18, Supplement, 1-53.
 - Fink A, Yano EM e Brook RH, 1989. The Condition of the Literature on Differences in Hospital Mortality. *Medical Care*, 27 (4), 315-336.
 - Flanders WD, Tucker G, Krishnadasan A, Martin D, Honig E e McClellan WM, 1999. Validation of Pneumonia Severity Index – Importance of Study-Specific Recalibration. *Journal of General Internal Medicine*, 14 (6), 333-340.
 - Forman DE, Berman AD, McCabe CH, Baim DS e Wey JY, 1992. PTCA in the Elderly: The Young-Old versus the Old-Old. *Journal of the American Geriatric Society*, 40 (1), 19-22.
 - Francks P, Gold MR e Clancy CM, 1996. Use of Care and Subsequent Mortality: The Importance of Gender. *Health Services Research*, 31 (3), 347-363.
 - Garg ML, Louis DZ, Glibe WA, Spirka CS, Skipper CS e Parekh RR, 1978. Evaluating Inpatient Costs: The Staging Mechanism. *Medical Care*, 16 (3), 191-201.
 - Garnick DW, DeLong ER e Luft HS, 1995. Measuring Hospital Mortality Rates: Are 30-Day Data Enough? *Health Services Research*, 29 (6), 679-695.
 - Gemmel P e Dierdonck RV, 1999. Admission Scheduling in Acute Care Hospitals: Does the Practice Fit with the Theory? *International Journal of Operations & Production Management*, 19 (9), 863-878.
 - Geraci JM, 2002. The Demise of Comparative Provider Complication Rates Derived from ICD-9-CM Code Diagnoses. *Medical Care*, 40 (10), 847-850.

-
- Gertman PM e Lowenstein S, 1984. A Research Paradigm for Severity of Illness: Issues for the Diagnosis-Related Group System. *Health Care Financing Administration, Annual Supplement*, 79-90.
 - Gijzen R, Hoeymans N, Schellevis FG, Ruwaard D, Satariano WA e Bos GA, 2001. Causes and Consequences of Comorbidity: A Review. *Journal of Clinical Epidemiology*, 54 (7), 661-674.
 - Given RS, 1996. Economies of Scale and Scope as an Explanation of Merger and Output Diversification Activities in the Health Maintenance Industry. *Journal of Health Economics*, 15 (6), 685-713.
 - Goldfarb MG, Hornbrook MC e Higgins CS, 1983. Determinants of Hospital Use: A Cross Diagnostic Analysis. *Medical Care*, 21 (1), 48-66.
 - Goldman L, Caldera DL, Nussbaum SR, Southwick FS, Krogstad D, Murray B, Burke DS, O'Malley TA, Goroll AH, Caplan CH, Nola J, Carabello B e Slater EE, 1977. Multifactorial Index of Cardiac Risk in Noncardiac Surgical Procedures. *New England Journal of Medicine*, 297 (16), 845-850.
 - Gonnella JS, Hornbrook MC e Louis DZ, 1984. Staging of Disease – A Case-Mix Measurement. *JAMA*, 251 (5), 637-644.
 - Gonnella JS, Louis DZ, Zeleznik C e Turner BJ, 1990. The Problem of Late Hospitalization: A Quality and Cost Issue. *Academic Medicine*, 65 (5), 314-319.
 - Gonnella JS, Louis DZ e Gozum ME, 1999 (editors). *Disease Staging – Clinical Criteria (Version 17)*. MEDSTAT Group, Santa Barbara, CA, 1999.
 - Gordon HS e Rosenthal GE, 1999. The Relationship of Gender and In-Hospital Death – Increased Risk of Death in Men. *Medical Care*, 37 (3), 318-324.
 - Gowrisankaran G e Town RJ, 2003. Competition, Payers, and Hospital Quality. *Health Services Research*, 38 (6), Part I, 1403-1421.
 - Green J, Wintfeld N e Passman LJ, 1990. The Importance of Severity of Illness in Assessing Hospital Mortality. *JAMA*, 263 (2), 241-246.
 - Green J, Passman LJ e Wintfeld N, 1991. Analyzing Hospital Mortality – The Consequences of Diversity in Patient Mix. *JAMA*, 265 (14), 1849-1853.
 - Green J e Wintfeld N, 1993. How Accurate Are Hospital Discharge Data for Evaluating Effectiveness of Care? *Medical Care*, 31 (8), 719-731.
 - Greenfield S, Blanco DM, Elashoff RM e Ganz PA, 1987. Patterns of Care Related to Age of Breast Cancer Patients. *JAMA*, 257 (20), 2766-2770.
 - Greenfield S, Aronow HU, Elashoff RM e Watanabe D, 1988. Flaws in Mortality Data – The Hazards of Ignoring Comorbid Disease. *JAMA*, 260 (15), 2253-2255.
 - Greenfield S, Apolone G, McNeil BG e Cleary PD, 1993. The Importance of Co-Existent Disease in the Occurrence of Post-Operative Complications and One-Year Recovery in Patients Undergoing Total Hip Replacement. *Medical Care*, 31 (2), 141-154.
 - Greenland S, 1984. Bias Methods for Deriving Standardized Morbidity Ratio and Attributable Fraction Estimates. *Statistics in Medicine*, 3, 131-141.
 - Griffith JR, Alexander JA e Jelinek RC, 2000. Measuring Comparative Hospital Performance. *Journal of Healthcare Management*, 47 (1), 41-56.
 - Gross PA, Braun BI, Kritchevsky SB e Simmons BP, 2000. Comparison of Clinical Indicators for Performance Measurement of Health Care Quality: A Cautionary Note. *British Journal of Clinical Governance*, 5 (4), 202-211.
 - Gustafson DH, Fryback DG, Rose JH, Yick V, Prokop CT, Detmer DE e Moore J, 1986. A Decision Theoretic Methodology for Severity Index Development. *Medical Decision Making*, 6 (1), 27-35.
 - Harrel FA, Lee KL, Califf RM, Pryor DB e Rosati RA, 1984. Regression Modelling Strategies for Improved Prognostic Prediction. *Statistics in Medicine*, 3, 143-152.
 - Harrel FE, Marcus SE, Layde PM, Broste SK, Cook EF, Wagner DP, Muhlbaier LH e Peck SL, 1990. Statistical Methods in SUPPORT. *Journal of Clinical Epidemiology*, 43, Supplement, 89S-98S.
 - Hartz AJ, Guse C, Sigmann P, Krakauer H, Goldman RS e Hagen TC, 1994. Severity of Illness Measures Derived From the Uniform Clinical Data Set (UCDSS). *Medical Care*, 32 (9), 881-901.

-
- HCIA, 1999. One Hundred Top Hospitals – Benchmarks for Success. HCIA, L.L.C., Evanston, Illinois, 1999.
 - Hill CA, Winfrey KL e Rudolph BA, 1997. “Best Hospitals”: A Description of the Methodology for the Index of Hospital Quality. *Inquiry*, 34 (1), 80-90.
 - Hofer TP e Hayward RA, 1996. Identifying Poor-Quality Hospitals – Can Hospital Mortality Rates Detect Quality Problems for Medical Diagnoses? *Medical Care*, 34 (8), 737-753.
 - Horn SD e Schumacher DN, 1979. An Analysis of Case Mix Complexity Using Information Theory and Diagnostic Related Grouping. *Medical Care*, 17 (4), 382-389.
 - Horn SD, Bulkey G, Sharkey PD, Chambers AF, Horn RA e Schramm CJ, 1985. Interhospital Differences in Severity of Illness – Problems for Prospective Payment based on Diagnosis Related Groups (DRGs). *The New England Journal of Medicine*, 313 (1), 20-24.
 - Horn SD, Horn RA e Moses H, 1986. Profiles of Physician Practice and Patient Severity of Illness. *American Journal of Public Health*, 76 (5), 532-535.
 - Horn SD, Sharkey PD, Bucle JM, Backofen JE, Averill RF e Horn RA, 1991. The Relationship Between Severity of Illness and Hospital Length of Stay and Mortality. *Medical Care*, 29 (4), 305-317.
 - Horn SD, 1986. Measuring Severity: How Sick Is Sick? How Well Is Well? *Healthcare Financing Management*, 40 (10) 21, 24-32.
 - Horn SD, 1988. Severity of Illness and the Adverse Patient Occurrence Index: A Reliability Study and Policy Implications. *Medical Care*, 26 (7), 736-738.
 - Horn SD, 1997. Physician Profiling. How Can it Be Misleading and What to Do. *Consultant*, 27 (12), 86-88 e 93-94.
 - Hornbrook MC e Goldfarb MG, 1983. A Partial Test of Hospital Behavioral Model. *Social Science & Medicine*, 17 (10), 667-680.
 - Hornbrook MC e Monheit AC, 1985. The Contribution of Case-Mix Severity to the Hospital Cost-Output Relation. *Inquiry*, 22 (3), 259-271.
 - Hornbrook MC, Goodman MG, 1996. Chronic Disease, Functional Health Status, and Demographics: A Multi-Dimensional Approach to Risk Adjustment. *Health Services Research*, 31 (3), 283-346.
 - Hornbrook MC, 1982. Hospital Case Mix: Its Definition Measurement and Use: Part I. The Conceptual Framework e Part II. Review of Alternative Measures. *Medical Care Review*, 39 (1), 1-43 e *Medical Care Review*, 39 (2), 73-123.
 - Hornbrook MC, 1985. Techniques for Assessing Hospital Case Mix. *Annual Review of Public Health*, 6, 295-324.
 - Hosmer DW, Jovanovic B e Lemeshow S, 1989. Best Subsets Logistic Regression. *Biometrics*, 45, 1265-1270.
 - Hosmer DW e Lemeshow S, 1989. *Applied Logistic Regression*. John Wiley & Sons, New York, 1989.
 - Houchens RL, 2002. A General Method for Measuring Inpatient Comorbidity Burden on Inpatient Mortality from a Large Administrative Database. *MEDSTAT*, Santa Barbara, CA, 2002.
 - Hughes JS, Iezzoni LI, Daley J e Greenberg L, 1996. How Severity Measures Rate Hospitalized Patients. *Journal of General Internal Medicine*, 11 (5), 303-311.
 - Ibrahim JE, Majoor JW, Boyce MW e McNeil JJ, 1998. Pilot Hospital-Wide Clinical Indicators Project – Final Report. Commonwealth of Australia, Canberra, 1998.
 - Iezzoni LI e Moskowitz MA, 1986. Clinical Overlap among Medical Diagnosis-Related Groups. *JAMA*, 255 (7), 927-929.
 - Iezzoni LI, Ash AA, Cobb JL e Moskowitz MA, 1988. Admission MedisGroups Score and the Cost of Hospitalization. *Medical Care*, 26 (11), 1068-1080.
 - Iezzoni LI, Shwartz M e Restuccia J, 1991. The Role of Severity Information in Health Policy Debates: A Survey of State and Regional Concerns. *Inquiry*, 28 (2), 117-128.

- Iezzoni LI, Ash AA, Coffman G e Moskowitz MA, 1991a. Admission and Mid-Stay MedisGroups Scores as Predictors of Death within 30 Days of Hospital Admission. *American Journal of Public Health*, 81 (1), 74-78.
- Iezzoni LI, Henderson MG, Bergman A e Drews RE, 1991b. Purpose of Admission and Resource Use During Cancer Hospitalizations. *Health Care Financing Review*, 13 (2), 29-40.
- Iezzoni LI e Daley J, 1992. A Description and Clinical Assessment of the Computerized Severity Index. *Quality Review Bulletin*, February, 44-52.
- Iezzoni LI, Ash AS, Coffman GA, Moskowitz MA, 1992a. Predicting In-Hospital Mortality – A Comparison of Severity Measurement Approaches. *Medical Care*, 30 (4), 347-359.
- Iezzoni LI, Restuccia JD, Shwartz M, Schaumburg D, Coffman GA, Kreger BE, Butterly JR e Selker HP, 1992b. The Utility of Severity of Illness Information in Assessing the Quality of Hospital Care – The Role of Clinical Trajectory. *Medical Care*, 30 (5), 428-444.
- Iezzoni LI, Foley SM, Heeren T, Daley J, Duncan C, Fisher ES e Hughes J, 1992c. A Method for Screening the Quality of Hospital Care Using Administrative Data: Preliminary Validation Results. *Quality Research Bulletin*, November, 361-371.
- Iezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES e Heeren T, 1992d. Comorbidities, Complications and Coding Bias – Does the Number of Diagnosis Codes Matter in Predicting In-Hospital Mortality? *JAMA*, 267 (16), 2197-2203.
- Iezzoni LI, Hotchin EK, Ash AA, Shwartz M e Mackiernan Y, 1993. MedisGroups Data Bases – The Impact of Data Collection Guidelines on Predicting In-Hospital Mortality. *Medical Care*, 31 (3), 277-283.
- Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J e Coffman GA, 1994a. Chronic Conditions and Risk of In-Hospital Death. *Health Services Research*, 29 (4), 435-460.
- Iezzoni LI, Daley J, Heeren T, Foley SM, Hughes JS, Fisher ES, Duncan CC e Coffman GA, 1994b. Using Administrative Data to Screen Hospitals for High Complications Rates. *Inquiry*, 31 (1), 40-55.
- Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, Hughes JS e Coffman GA, 1994c. Identifying Complications of Care Using Administrative Data. *Medical Care*, 32 (7), 700-715.
- Iezzoni LI, Ash AA, Shwartz M, Daley J, Hughes J e Mackiernan YD, 1995a. Predicting Who Dies Depends on How Severity is Measured: Implications for Evaluating Patient Outcomes. *Annals of Internal Medicine*, 123 (10), 763-770.
- Iezzoni LI, Shwartz M, Ash AA, Hughes JS, Daley J e Mackiernan YD, 1995b. Using Severity-Adjusted Stroke Mortality Rates to Judge Hospitals. *International Journal for Quality in Health Care*, 7 (2), 81-94.
- Iezzoni LI, Shwartz M, Ash AA e Mackiernan YD, 1996a. Using Severity Measures to Predict the Likelihood of Death for Pneumonia Inpatients. *Journal of General Internal Medicine*, 11 (1), 23-31.
- Iezzoni LI, Shwartz M, Ash AA, Hughes JS, Daley J e Mackiernan YD, 1996b. Severity Measurement Methods and Judging Hospital Death Rates for Pneumonia. *Medical Care* 34 (1), 11-28.
- Iezzoni LI, Shwartz M, Ash AA e Mackiernan YD, 1996c. Does Severity Explain Differences in Hospital Length of Stay for Pneumonia Patients? *Journal of Health Services Research*, 1 (2), 65-76.
- Iezzoni LI, Shwartz M, Ash AA e Mackiernan YD, 1996d. Predicting In-Hospital Mortality for Stroke Patients: Results Differ across Severity-Measurement Methods. *Medical Decision Making* 16 (4), 248-256.
- Iezzoni LI, Ash AA, Shwartz M, Daley J, Hughes J e Mackiernan YD, 1996e. Judging Hospitals by Severity-Adjusted Mortality Rates: The Influence of the Severity-Adjustment Method. *American Journal of Public Health*, 86 (10), 1379-1387.
- Iezzoni LI, Ash AA, Shwartz M e Mackiernan YD, 1997. Differences in Procedure Use, In-Hospital Mortality, and Severity by Gender for Acute Myocardial Infarction Patients – Are Answers Affected by Data Source and Severity Measure? *Medical Care*, 35 (2), 158-171.
- Iezzoni LI, Mackiernan YD, Cahalane MJ, Philips RS, Davis RB e Miller K, 1999. Screening Inpatient Quality Using Post-Discharge Events. *Medical Care*, 37 (4), 384-398.

-
- Iezzoni LI, 1989. Using Severity Information for Quality Assessment: A Review of Three Cases by Five Severity Measures. *Quality Research Bulletin*, December, 376-382.
 - Iezzoni LI, 1990. Using Administrative Diagnostic Data to Assess the Quality of Hospital Care. *International Journal of Technology Assessment in Health Care*, 6 (2), 272-281.
 - Iezzoni LI, 1993. Monitoring Quality of Care: What Do We Need to Know? *Inquiry*, 30 (2), 112-114.
 - Iezzoni LI, 1995. Risk Adjustment for Medical Effectiveness Research: An Overview of Conceptual and Methodological Considerations. *Journal of Investigate Medicine*, 43 (2), 136-150.
 - Iezzoni LI, 1996a. An Introduction to Risk Adjustment. *American Journal of Medical Quality*, 11 (1), S8-S11.
 - Iezzoni LI, 1996b. On Opening "Black Boxes" and Looking Inside. *International Journal of Quality in Health Care*, 8 (3), 209-210.
 - Iezzoni LI, 1996c. 100 Apples Divided by 15 Red Herrings: A Cautionary Tale From the Mid-19th Century on Comparing Hospital Mortality Rates. *Annals of Internal Medicine*, 24 (12), 1079-1085.
 - Iezzoni LI, 1997a. Case Classification and Quality of Care: Issues to Consider Before Making the Investment. *Quality Research Bulletin*, April, 135-139.
 - Iezzoni LI, 1997b. Risks and Outcomes. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 1-40.
 - Iezzoni LI, 1997c. The Risks of Risk Adjustment. *JAMA*, 278 (19), 1600-1607.
 - Iezzoni LI, 1997d. Dimensions of Risk. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 43-167.
 - Iezzoni LI, 1997e. Assessing Quality Using Administrative Data. *Annals of Internal Medicine*, 127 (8, Part 2), 666-674.
 - Iezzoni LI, 1997f. Data Sources and Implications: Administrative Databases. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 169-242.
 - Iezzoni LI, 1999. Statistically Derived Predictive Models – Caveat Emptor. *Journal of General Internal Medicine*, 14 (6), 388-389.
 - Instituto Nacional de Administração, 1999. Avaliação dos Hospitais Fernando da Fonseca e Garcia de Orta. INA, Lisboa, 1999.
 - Instituto Nacional de Administração, 2001. Projecto de Avaliação de Unidades de Saúde. INA, Lisboa, 2001.
 - Jacobs P, 1974. A Survey of Economic Models of Hospitals. *Inquiry* 11 (2), 83-97.
 - Jencks SF e Dobson A, 1987. Refining Case-Mix Adjustment. The Research Evidence. *New England Journal of Medicine*, 317 (11), 679-686.
 - Jencks SF, Daley J, Draper D, Thomas N, Lenhart G e Walker J, 1988. Interpreting Hospital Mortality Data – The Role of Clinical Risk Adjustment. *JAMA*, 260 (24), 3611-3616.
 - Jencks SF, Dobson A, Willis, P e Feinstein PH, 1984. Evaluating and Improving the Measurement of Hospital Case Mix. *Health Care Financing Administration, Annual Supplement*, 1-11.
 - Jencks SF, Williams DK e Kay TL, 1988. Assessing Hospital-Associated Deaths from Discharge Data. The Role of Length of Stay and Comorbidities. *JAMA*, 260 (15), 2240-2246.
 - Johnson ML, 1996. Evaluating the Performance of Risk Adjustment Methods. *Outcomes Assessment Research*, School of Public Health of University of Texas, Houston, Texas, 1996.
 - Jones KR, 1985. Predicting Hospital Charge and Stay Variation – The Role of Patient Teaching Status, Controlling for Diagnosis Related Groups, Demographic Characteristics, and Severity of Illness. *Medical Care*, 23 (3), 220-235.
 - Justice AC, Covinsky KE e Berlin JA, 1999. Assessing the Generability of Prognostic Information. *Annals of Internal Medicine*, 130 (5), 515-524.

-
- Kaboli PJ, Barnett MJ, Fuehrer SM e Rosenthal GE, 2001. Length of Stay as a Source of Bias in Comparing Performance in VA and Private Sector Facilities. *Medical Care*, 39 (9), 1014-1024.
 - Kalish RL, Daley J, Duncan C, Davis RB, Coffma GA e Iezzoni LI, 1995. Costs of Potential Complications of Care for Major Surgery Patients. *American Journal of Medical Quality*, 10 (1), 48-54.
 - Kanouse DE, Kallich JD e Kahan JP, 1995. Dissemination of Effectiveness and Outcomes Research. *Health Policy*, 34 (3), 167-192.
 - Kaplan SH, Greenfield S e Ware JR, 1989. Assessing the Effects of Physician-Patient Interactions on the Outcomes of a Chronic Disease. *Medical Care*, 27 (3), supplement, S110-S127.
 - Kazis LE, 1991. Health Outcome Assessments in Medicine: History, Applications, and New Directions. *Advanced Internal Medicine*, 36, 109-130.
 - Kelleher C, 1993. Relationship of Physician Ratings of Severity and Difficulty of Clinical Management to Length of Stay. *Health Services Research*, 27 (6), 841-855.
 - Klastorin TD e Watts CA, 1982. A Current Reappraisal of Berry's Hospital Typology. *Medical Care*, 20 (5), 441-449.
 - Knaus WA, Zimmerman JE, Wagner DP e Lawrence DE, 1981. APACHE – Acute Physiology and Chronic Health Evaluation: A Physiologically Based Classification System. *Critical Care Medicine*, 9 (8), 591-597.
 - Knaus WA, Wagner DP e Draper EA, 1984. The Value of Measuring Severity of Disease in Clinical Research on Acutely Ill Patients. *Journal of Chronic Diseases*, 37 (6), 455-463.
 - Knaus WA, Draper EA, Wagner DP e Zimmerman JE, 1985. APACHE II: A Severity of Disease Classification System. *Critical Care Medicine*, 9 (8), 591-597.
 - Knaus WA, Draper EA, Wagner DP e Zimmerman JE, 1986. An Evaluation of Outcome from Intensive Care in Major Medical Centers. *Annals of Internal Medicine*, 104 (3), 410-418.
 - Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A e Harrel FE, 1991. The APACHE III Prognostic System: Risk Prediction of Mortality for Critically Ill Hospitalized Adults. *Chest*, 100 (6), 1619-1636.
 - Knaus WA, Wagner DP, Zimmerman JE e Draper EA, 1993. Variations in Mortality and Length of Stay in Intensive Care Units. *Annals of Internal Medicine*, 118 (10), 753-761.
 - Krakauer H e Jacoby I, 1993. Predicting the Course of Disease. *Inquiry*, 30 (2), 115-127.
 - Krakauer H, Bailey RC, Skellan KJ, Stewart JD, Hartz AJ, Kuhn EM e Rimm AA, 1992. Evaluation of the HCFA Model for the Analysis of Mortality Following Hospitalization. *Health Services Research*, 27 (3), 317-335.
 - Kulinskaya E, Staudte RG e Hales C, 2002. Does Weighted ANOVA and Weighted R2 Work on Real Length of Stay Data? 18th PCS/E Conference, Innsbruck, Austria, 2002.
 - Landis JR e Koch GG, 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33 (1), 159-174.
 - Landon B, Iezzoni LI, Ash AA, Shwartz M, Daley J, Hugues JS e Mackiernan YD, 1996. Judging Hospitals by Severity-Adjusted Mortality Rates: The Case of CABG Surgery. *Inquiry*, 33 (2), 155-166.
 - Lave JR e Leinhardt S, 1976. An Evaluation of a Hospital Stay Regulatory Mechanism. *Inquiry*, 13 (4), 327-342.
 - Lave JR e Franck RG, 1990. Effect of the Structure of Hospital Payment on Length of Stay. *Health Services Research*, 25 (2), 327-347.
 - Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH e Iezzoni LI, 2000. Identification of In-Hospital Complications From Claims Data – Is It Valid? *Medical Care*, 38 (8), 785-795.
 - Le Gall JR, Lemeshow S e Saulnier F, 1993. A New Simplified Acute Physiology Score (SAPSII) Based on a European/North American Multicenter Study. *JAMA*, 270 (24), 2957-2963.
 - Lemeshow S e Hosmer DW, 1982. A Review of Goodness of Fit Statistics for Use in the Development of Logistic Regression Models. *American Journal of Epidemiology*, 115 (1), 92-106.

-
- Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH e Rapoport J, 1993. Mortality Probability Models (MPM II) Based on an International Cohort of Intensive Care Units Patients. *JAMA*, 270 (20), 2478-2486.
 - Lichtig LK, 1986. Hospital Information Systems for Case Mix Management. John Wiley & Sons, New York, 1986.
 - Lohr KN, 1988. Outcome Measurement: Concepts and Questions. *Inquiry* 25 (1), 37-50.
 - Lohr KN, 1990. Use of Insurance Claims Data in Measuring Hospital Quality of Care. *International Journal of Technology Assessment*, 6 (2), 263-271.
 - Louis DZ, 2003. Potentially Inappropriate Ordinary Hospital Admissions in the Regione Emilia-Romagna. The Center for Research in Medical Education and Health Care, Thomas Jefferson University, Philadelphia, 2003.
 - Luke RD, 1979. Dimensions in Hospital Case Mix Measurement. *Inquiry*, 16 (1), 38-49.
 - Manheim LM, Feinglass J, Shortell SM e Hughes EF, 1992. Regional Variation in Medicare Hospital Mortality. *Inquiry*, 29 (1), 55-66.
 - Marazzi A, Paccaud F, Ruffieux C e Begun C, 1998. Fitting the Distributions of Length of Stay by Parametric Models. *Medical Care*, 36 (6), 915-927.
 - Markson LE, Nash DB, Louis DZ e Gonnella JS, 1991. Clinical Outcomes Management and Disease Staging. *Evaluation & The Health Professions*, 14 (2), 201-227.
 - Marquis SM e Buchanan JL, 1999. Simulating the Effects of Employer Contributions on Adverse Selection and Health Plan Choice. *Health Services Research*, 34 (4), 813-837.
 - McCarthy EP, Iezzoni LI, Davis RB, Palmer TH, Cahalane M, Hamel MB, Mukamal K, Philips RS e Davies DT, 2000. Does Clinical Evidence Support ICD-9-CM Diagnosis Coding of Complications? *Medical Care*, 38 (8), 868-876.
 - McGuire A e Williams H, 1986. Information Theory and Scottish Hospital Cost Functions. University of Aberdeen, Discussion paper N° 01/86, Aberdeen, 1986.
 - McGuire A, Henderson J e Mooney G, 1988. The Economics of Health Care – An Introductory Text. Routledge, London, 1988.
 - McGuire A, 1989. The Measurement of Hospital Efficiency: a Reply. *Social Science & Medicine*, 28 (1), 13-17.
 - McGuire TE, 1991. An Evaluation of Diagnosis Related Group Severity and Complexity Refinement. *Health Care Financing Review*, 12 (4), 49-60.
 - McMahon LF e Billi JE, 1988. Measurement of Severity of Illness and the Medicare Prospective Payment System: State of the Art and Future Directions. *Journal of General Internal Medicine*, 3 (5), 482-490.
 - McMahon LF e Smits HL, 1986. Can Medicare Prospective Payment Survive the ICD-9-CM Disease Classification System? *Annals of Internal Medicine*, 104 (4), 562-566.
 - McNeil BK, Haley JA, 1984. Statistical Approaches to the Analysis of Receiver Operating Characteristic (ROC) Curves. *Medical Decision Making*, 4 (2), 137-150.
 - MEDSTAT, 2001. Disease Staging Software, Version 4.12 – User Guide. The MEDSTAT Group, Inc., Ann Arbor, Michigan, 2001.
 - MEDSTAT, 2002. Disease Staging – Calibration and Recalibration Procedures. The MEDSTAT Group, Inc., Ann Arbor, Michigan, 2001.
 - Meenan RT, Goodman MJ, Fishman PA, Hornbrook MC, Rosetti MC e Bachman DJ, 2003. Using Risk-Adjustment Models to Identify High-Cost Risks. *Medical Care*, 41 (11), 1301-1312.
 - Melo FG, 1985. Métodos Estatísticos em Estudos Comparativos: Comparação de Tratamentos. Escola Nacional de Saúde Pública, Lisboa, 1985.
 - Menemeyer ST, Morrissey MA e Howard, LZ, 1997. Death and Reputation: How Consumers Acted Upon HCFA Mortality Information. *Inquiry*, 34 (2), 117-128.

-
- Miguel JP e Costa C, 1997. A Reforma da Saúde em Portugal: À Procura da Eficiência. *Revista Portuguesa de Saúde Pública*, 15 (2), 5-17.
 - Ministério da Saúde, 2001. Lei nº27/2002 de 8 de Novembro (Novo regime Jurídico da Gestão Hospitalar). *Diário da República*, 258, I Série – A.
 - Ministério da Saúde, 2001. Portaria nº189/2001 de 9 de Março. *Diário da República*, 58, I Série – B.
 - Moreno R, Apolone G e Miranda DR, 1998. Evaluation of the Uniformity of Fit of General Outcome Prediction Models. *Intensive Care Medicine*, 24 (1), 40-47.
 - Mozes B, Easterling MJ, Sheiner LB, Melmon KL, Kline R, Goldman ES e Brown AN, 1994. Case-Mix Adjustment Using Objective Measures of Severity: The Case for Laboratory Data. *Health Services Research*, 28 (6), 689-712.
 - Mullin RL, 1985. Diagnosis-Related Groups and Severity. ICD-9-CM, the Real Problem. *JAMA*, 254 (9), 1208-1210.
 - Mullin RL, Averill RF e Boucher SA, 2002. A Comparison of American and Australian DRG Systems. 18th PCS/E, Innsbruck, Austria, 2002.
 - Murphy DJ e Cluff LE (editors), 1990. SUPPORT: Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. Study Design. *Journal of Clinical Epidemiology*, 43 (supplement).
 - Naessens JM, Leibson CL, Krishan I e Ballard DJ, 1992. Contribution of a Measure of Disease Complexity (COMPLEX) to Prediction of Outcome and Charges Among Hospitalized Patients. *Mayo Clinic Process*, 67, 1140-1149.
 - NHPC, 2001. National Health Performance Framework Report. National Health Performance Committee, Queensland Health, Brisbane, 2001.
 - NHS, 2002. NHS Performance Indicators – Acute NHS Hospital Trusts: February, 2002. National Health Service, United Kingdom, 2002.
 - Nóbrega SD, 2004. A Visão Estratégica para a Rede SA. Unidade de Missão Hospitais SA, Lisboa, 2004.
 - O'Muircheartaigh C, Murphy J e Moore W, 2002. The 2002 Index of Hospital Quality. NORC, University of Chicago, Chicago, 2002.
 - Paiva RL, 1993. Eficiência Técnica e Eficiência de Afecção no Sistema Hospitalar Português. III Encontro da Associação Portuguesa de Economia da Saúde – Comunicações Apresentadas. APES, Lisboa, 1993.
 - Pereira J, 1993. Economia da Saúde – Glossário de Termos e de Conceitos. Lisboa, Associação Portuguesa de Economia da Saúde, Documento de Trabalho 1/93, Lisboa, 1993.
 - Pettengill J e Vertrees J, 1980. New Uses for Old Data: A Medicare Hospital Case Mix Index. Proceedings of the 18th National Meeting of the Public Health Conference on Records and Statistics, Washington, 1980.
 - Pettengill J e Vertrees J, 1982. Issues of Reliability and Validity in Hospital Case Mix Measurement. *Health Care Financing Review*, 4(2), 101-128.
 - Philips RS e Knaus WA, 1990. Patient Characteristics in SUPPORT: Sociodemographics, Admission Diagnosis, Co-morbidities and Acute Physiology Score. *Journal of Clinical Epidemiology*, 43, Supplement, 29S-31S.
 - Philips RS, Hamel MB, Covinsky KE e Lynn J, 2000 (editors). Findings from SUPPORT and HELP. *Journal of the American Geriatrics Society*, 48, Supplement.
 - Plomman MP, 1985. Choosing a Patient Classification System to Describe the Hospital Product. *Hospital & Health Services Administration*, 30 (3), 106-117.
 - Polanczyck CA, Rhode LE, Philbin EA e Salvo TG, 1998. A New Casemix Adjustment Index for Hospital Mortality Among Patients With Congestive Heart Failure. *Medical Care*, 36 (10), 1489-1499.
 - Polverejan E, Gardiner JC, Bradley CJ e Rovner MH, 2003. Estimation of Mean Hospital Cost as a Function of Length of Stay and Patient Characteristics. *Health Economics*, 12 (11), 935-947.

- Pollack MM, Ruttiman UE e Getson PR, 1988. Pediatric Risk of Mortality (PRISM) Score. *Critical Care Medicine*, 16 (11), 1110-1116.
- Pollack MM, Patel KM e Ruttiman UE, 1996. PRISM III: An Updated Pediatric Risk of Mortality. *Critical Care Medicine*, 24 (5), 743-752.
- Powel H, Lim LL e Heller RF, 2001. Accuracy of Administrative Data to Assess Comorbidity in Patients with Heart Disease: An Australian Perspective. *Journal of Clinical Epidemiology*, 54, 687-693.
- Pryor DB e Lee KL, 1991. Methods for the Analysis and Assessment of Clinical Databases: The Clinician's Perspective. *Statistics in Medicine*, 10, 617-628.
- Quality Measurement Advisory Service, 1997. *Quality Measurement Tools – Assessing Hospital Performance*. QMAS, Washington, 1997.
- Quan H, Parsons GA e Ghali WA, 2002. Validity of Information on Comorbidity Derived From ICD-9-CM Administrative Data. *Medical Care*, 40 (8), 675-685.
- Rapoport J, Teres D, Zhao Y e Lemeshow S, 2003. Length of Stay as a Guide to Hospital Economic Performance for ICU Patients. *Medical Care*, 41 (3), 386-397.
- Ribeiro JM, 2004. *Hospitais SA: Balanço do Primeiro Ano de Empresarialização*. Unidade de Missão Hospitais SA, Lisboa, 2004.
- Robinson JC e Gardner LB, 1995. Adverse Selection Among Multiple Competing Health Maintenance Organizations. *Medical Care*, 33 (12), 1161-1175.
- Rodrigues S, 2003. *Análise da Produção do Internamento do Hospital de São Teotónio – Viseu: Os GDHs enquanto Instrumento de Gestão*. Dissertação do XXXI Curso de Especialização em Administração Hospitalar, Escola Nacional de Saúde Pública, Lisboa, 2003.
- Romano PS, Chan BJ, Schembri ME e Rainwater JA, 2002. Can Administrative Data Be Used to Compare Postoperative Complication Rates Across Hospitals? *Medical Care*, 40 (10), 856-867.
- Rosen HM e Gree BA, 1987. The HCFA Excess Mortality Lists: A Methodological Critique. *Hospital & Health Services Administration*, February, 119-127.
- Rosko MD, 1988. DRGs and Severity of Illness Measures: An Analysis of Patient Classification Systems. *Journal of Medical Systems*, 12 (4), 257-274.
- Rubenstein LV, Calkins DR, Young RT, Cleary PD, Fink A, Kosekoff J, Jette AM, Davies AR, Delbano TL e Brook RH, 1989. Improving Patient Function: A Randomized Trial of Functional Disability Screening. *Annals of Internal Medicine*, 111 (10), 836-842.
- Ruttiman UE, 1994. Statistical Approaches to Development and Validation of Predictive Instruments. *Critical Care Clinics*, 10 (1), 19-35.
- Santana R, 2003. *O Financiamento Hospitalar e a Definição de Preços*. Dissertação do XXXI Curso de Especialização em Administração Hospitalar, Escola Nacional de Saúde Pública, Lisboa, 2003.
- Scitovsky AA, McNall N e Benham L, 1978. Factors Affecting the Choice between Two Prepaid Plans. *Medical Care*, 16 (8), 660-681.
- Schumacher DN, Parker B, Kofie V e Munns JM, 1987. Severity of Illness Index and the Adverse Patient Occurrence Index: A Reliability Study and Policy Implications. *Medical Care*, 25 (8), 695-704.
- Schumacher DN, Horn SD, Solnick MF, Atkinson G e Cook J, 1979. Hospital Cost per Case: Analysis Using a Statewide Data System. *Medical Care*, 17 (10), 1037-1047.
- Schweiker RS, 1982. *Report to Congress: Hospital Prospective Payment for Medicare*. Health Care Financing Administration, Washington, DC, 1982.
- Selim AJ, Berlowitz DR, Fincke G, Rosen A, Ren XS, Christiansen CL, Cong Z, Lee A e Kazis L, 2002. Risk-Adjusted Mortality Rates as a Potential Outcome Indicator for Outpatient Quality Assessments. *Medical Care*, 40 (3), 237-245.
- Serrão D, Abrantes A, Veloso AB, Oliveira G, Moreira JM, Delgado M e Sousa MD, 1998. *Reflexão Sobre a Saúde – Recomendações para Uma Reforma Estrutural*. Conselho de Reflexão sobre a Saúde, Lisboa, 1998.

-
- Shapiro JR, Dykstra DM, Pisoni R, Beronia N, Gaylin DS, Oppenheimer CC, Rubin RJ e Helad PJ, 2003. Patient Selection in the ESRD Managed Care Demonstration. *Health Care Financing Review*, 24 (4), 31-43.
 - Shen Y e Ellis RP, 2002. How Profitable is Risk Selection? A Comparison of Four Risk Adjustment Models. *Health Economics*, 11 (2), 165-174.
 - Shen YC, 2003. The Effect of Financial Pressure on the Quality of Care in Hospitals. *Journal of Health Economics*, 22 (2), 243-269.
 - Shwartz M e Ash AA, 1997. Evaluating the Performance of Risk Adjustment Methods. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 427-469.
 - Shwartz M, Ash AA e Iezzoni LI, 1997. Comparing Outcomes Across Providers. In Iezzoni LI (editor), "Risk Adjustment for Measuring Healthcare Outcomes", 2nd Edition, Health Administration Press, Chicago, 1997, 471-516.
 - Shwartz M, Iezzoni LI, Ash AA e MacKiernan YD, 1996. Do Severity Measures Explain Differences in Length of Hospital Stay? The Case of Hip Fracture. *Health Services Research*, 31 (4), 365-385.
 - Simborg DW, 1981. DRG Creep: A New Hospital-Acquired Disease. *New England Journal of Medicine*, 304 (26), 1602-1604.
 - Smith PJ, Thompson TJ, Engelgau MM e Herman WH, 1996. A Generalized Linear Model for Analyzing Receiver Operator Characteristic Curves. *Statistics in Medicine*, 15, 323-333.
 - Snelling I, 2003. Do Star Ratings Really Reflect Hospital Performance? *Journal of Health Organization and Management*, 17 (3), 210-223.
 - Stearns SC, 1991. Hospital Discharge Decisions, Health Outcomes, and the Use of Unobserved Information on Case-Mix Severity. *Health Services Research*, 26 (1), 27-51.
 - Steen PM, Brewster AC, Bradbury RC, Estabrook E e Young JA, 1993. Predicted Probabilities of Hospital Death as a Measure of Admission Severity of Illness. *Inquiry*, 30 (2), 128-141.
 - Taroni F, Louis DZ, Yuen EJ, Anemona A e Zappi A, 1991. Timeliness of Hospital Admissions. 7th International PCS/E, Lausanne, Switzerland, 1991.
 - Taroni F, Repetto F, Louis DZ, Moro ML, Yuen EJ, Gonnella JS, 1987. Variation in Hospital Use and Avoidable Patient Morbidity. *Journal of Health Services Research and Policy*, 2 (4), 217-222.
 - Tatchell M, 1983. Measuring Hospital Output: A Review of the Service Mix and Case Mix Approaches. *Social Science & Medicine*, 17 (13), 871-883.
 - Thomas JW, Ashcraft ML e Zimmerman J, 1986. An Evaluation of Alternative Severity of Illness Measures for Use by University Hospitals. Department of Health Services Management and Policy, The University of Michigan, Ann Arbor, Michigan, 1986.
 - Thomas JW e Ashcraft ML, 1989. Measuring Severity of Illness: a Comparison of Interrater Reliability among Severity Methodologies. *Inquiry*, 26 (4), 483-492.
 - Thomas JW e Longo DR, 1990. Application of Severity Measurement Systems for Hospital Quality Measurement. *Hospital & Health Services Administration*, 35 (2), 221-243.
 - Thomas JW e Ashcraft ML, 1991. Measuring Severity of Illness: Six Severity Systems and Their Ability to Explain Cost Variations. *Inquiry*, 28 (1), 39-55.
 - Thomas JW, Holloway JJ e Guire KE, 1993. Validating Risk-Adjusted Mortality as an Indicator for Quality of Care. *Inquiry*, 30 (1), 6-22.
 - Thomas JW, Bates EW, Hofer T, Perkins A, Murphy NF e Webb C, 1998. Interpreting Risk-Adjusted Length of Stay Patterns for VA Hospitals. *Medical Care*, 36 (12), 1660-1675.
 - Thomas JW e Hofer TP, 1998. Research Evidence on the Validity of Risk-Adjusted Mortality Rate as a Measure of Hospital Quality of Care. *Medical Care Research and Review*, 55 (4), 371-404.
 - Thomas JW e Hofer TP, 1999. Accuracy of Risk-Adjusted Mortality Rate as a Measure of Hospital Quality of Care. *Medical Care*, 37 (1), 83-92.

-
- Thomas JW, 1996. Does Risk-Adjusted Readmission Rate Provide Valid Information on Hospital Quality? *Inquiry*, 33(3), 258-270.
 - Tully L e Rulon V, 2000. Evolution of the Uses of ICD-9-CM Coding: Medicare Risk Adjustment Methodology for Managed Care Plans. *Topics in Health Information Management*, 21 (2), 62-67.
 - Urbano J e Bentes M, 1990. Definição da Produção Hospitalar: Os Grupos de Diagnósticos Homogêneos. *Revista Portuguesa de Saúde Pública*, 8 (1), 49-60.
 - Vladeck BC, Goodwin EJ, Myers LP e Sinisi M, 1988. Consumers and Hospital Use: the HCFA "death list". *Health Affairs*, 7 (1), 122-125.
 - Vladeck BC e Kramer PS, 1988. Case Mix Measures: DRGs and Alternatives. *Annual Review of Public Health*, 9, 333-359.
 - Vladeck BC, 1984. Medicare Hospital Payment by Diagnosis-Related Groups. *Annals of Internal Medicine*, 100 (4), 576-591.
 - Voss BW, Hasman A, Rutten F, Zwaan C e Carpay JJ, 1994. Explaining Cost Variations in DRGs "Acute Myocardial Infarction" by Severity of Illness. *Health Policy*, 28 (1), 37-50.
 - Wagner DP, Knaus WA e Draper EA, 1983. Statistical Validation of a Severity of Disease Measure. *American Journal of Public Health*, 73 (8), 878-884.
 - Wagner DP, Knaus WA e Bergner M, 1989. Statistical Methods. *Critical Care Medicine*, 17 (12), Supplement, S194-S198.
 - Wagner DP, Knaus WA, Harrel FA, Zimmerman JE e Watts C, 1994. Daily Prognostic Estimates for Critically Ill Adults in Intensive Care Units: Results from a Prospective, Multicenter, Inception Cohort Analysis. *Critical Care Medicine*, 22 (9), 1359-1372.
 - Wagstaff, A, 1999. Econometric Studies in Health Economics – A Survey of the British Literature. Centre for Health Economics, University of York, 407 (8), York, 1999.
 - Walley T, 1999. Outcomes, Outcomes Research and Disease Management – A View from the UK. *Disease Management & Health Outcomes*, 5 (4), 197-207.
 - Watts CA e Klastorin TD, 1980. The Impact of Case Mix on Hospital Cost: A Comparative Analysis. *Inquiry*, 17 (4), 357-367.
 - Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, Philips RS, Davies DT e Banks NJ, 2000. Use of Administrative Data to Find Substandard Care: Validation of the Complications Screening Program. *Medical Care*, 38 (8), 796-806.
 - Weingart SN, Mukamal K, Davis RB, Davies Jr DT, Palmer RH, Chalane M, Hamel MB, Philips RS e Iezzoni LI, 2001. Physician-Reviewers' Perceptions and Judgements about Quality of Care. *International Journal for Quality in Health Care*, 13 (5), 357-365.
 - Weintraub WS, Craver JM, Cohen CL, Jones EL e Guyton RA, 1991. Influence of Age on Results of Coronary Artery Surgery. *Circulation*, 84 (5), Supplement III, III226-III235.
 - Westert GP, Nieboer AP e Groenewegen PP, 1993. Variation in Duration of Hospital Stay Between Hospitals and Between Doctors Within Hospitals. *Social Science & Medicine*, 37 (6), 833-839.
 - Wholey D, Feldman R, Christianson JB e Engberg J, 1996. Scale and Scope Economies among Health Maintenance Organizations. *Journal of Health Economics*, 15 (6), 657-684.
 - Wood WR, Ament RP e Kobrinsky EJ, 1981. A Foundation for Hospital Case Mix Measurement. *Inquiry*, 18 (3), 247-254.
 - Wray N, Hollingsworth JC, Petersen NJ e Ashton CM, 1997. Case-Mix Adjustment Using Administrative Databases: A Paradigm to Guide Future Research. *Medical Care Research Review*, 54 (3), 326-356.
 - Young WW, Kohler S e Kowalski J, 1994. PMC Patient Severity Scale: Derivation and Validation. *Health Services Research*, 29 (3), 367-390.
 - Young WW, Swinkola RB e Hutton MA, 1980. Assessment of the AUTOGRP Patient Classification System. *Medical Care*, 18 (2), 228-244.
 - Young WW, Swinkola RB e Zorn DM, 1982. The Measurement of Hospital Case Mix. *Medical Care*, 20 (5), 501-512.

- Young WW, 1984. Incorporating Severity of Illness and Comorbidity in Case-Mix Measurement. Health Care Financing Review, November, Annual Supplement, 23-31.
- Yuan Z, Cooper GS, Einstadter D, Cebul RD e Rimm AA, 2000. The Association Between Hospital Type and Mortality and Length of Stay. Medical Care, 38 (2), 231-245.
- Zalkind DL e Estaugh SR, 1997. Mortality Rates as an Indicator of Hospital Quality. Hospital & Health Services Administration, 42 (1), 3-15.
- Zimmerman JE, Shortell SM, Knaus WA, Rousseau DM, Wagner DP, Gillies RR, Draper EA e Devers K, 1993. Value and Cost of Teaching Hospitals: A Prospective, Multicenter, Inception Cohort Study. Critical Care Medicine, 21 (10), 1432-1442.
- Zimmerman JE, 1999 (editor). APACHE III Study Design, Analytic Plan of Evaluation of Severity and Outcome. Critical Care and Medicine, 17 (12), Part 2/Supplement.

ANEXOS

- Anexo I – Disease Staging
- Anexo II – Diagnosis Related Groups
- Anexo III – Demora Média – Outliers
- Anexo IV – Índices de Casemix
- Anexo V – Concentração da Produção
- Anexo VI – Complexidade dos Casos por Hospital
- Anexo VII – Gravidade dos Casos por Hospital
- Anexo VIII – Complexidade e Gravidade dos Casos por Hospital
- Anexo IX – Complexidade e Gravidade dos Casos por Doença (GADs)
- Anexo X – Efectividade
- Anexo XI – Eficiência
- Anexo XII – Desempenho Global

ANEXO I – DISEASE STAGING

- Doenças (exemplos)
- Estádios e Sub-Estádios (exemplos)
- Grandes Agrupamentos de Doenças (GADs)

Disease Staging

Doenças (exemplos)

Doenças do Sistema Nervoso Central (a) e (b)

Doenças – Códigos	Doenças – Designação
CN01	Meningite bacteriana
CN02	Meningite viral, encefalite e mielite
CN03	Síndrome de Reye
CN04	Tétano
CN05	Poliomielite
CN06	Toxoplasmose
CN07	Traumatismo da coluna e da espinal medula
CN08	Traumatismo craniano
CN09	Neoplasia do sistema nervoso central
CN10	Lesões do cérebro e da espinal medula
CN11	Demência degenera primária (doença de Alzheimer ou doença de Pick)
CN12	Coreia de Huntington
CN13	Esclerose Múltipla
CN14	Enxaqueca e cefaleias de tensão muscular
CN15	Atraso Mental
CN16	Epilepsia
CN17	Doença cerebrovascular
CN18	Traumatismos incisivos ou contusos da cabeça e pescoço, excepto os intracranianos
CN19	Lesões relacionadas com tratamentos do sistema nervoso central
CN70	Lesões cerebrais, excepto degenerescência cerebral, não classificadas noutros sítios
CN71	Síndromes paráliticos não classificados noutros sítios
CN72	Alterações do sistema nervoso central não classificadas noutros sítios
CN73	Degenerescência cerebral não classificada noutro sítio
CN74	Paralisia Cerebral
CN90	Alterações não específicas do cérebro
CN91	Paralisias não específicas e alterações do sistema nervoso central

Fonte: Gonnella, Louis e Gozum, 1999

(a) PDXCAT – no caso de Doença Principal e DXCAT no caso de Doenças Secundárias

Disease Staging

Estadios e Sub-Estadios (exemplos)

Estádios e Sub-Estádios (exemplos)

CN17 – Doença cerebrovascular

Estádio e Sub-Estádio	Doença	Código da CID-9-MC
1.1	Aterosclerose cerebral ou aneurisma cerebral	437.00; 437.10; 437.30; 437.40; 437.80; 437.90; 438.00; 442.81; 442.82; 483.89 e 438.90
2.1	Acidente isquémico transitório	433.00; 433.10; 433.20; 433.30; 433.80; 433.90 e 435.00-435.90
3.1	Hemorragia ou enfarte intracraniano ou intracerebral, trombose não piogénica dos seios venosos intracranianos ou doença moyamoya ou acidente vascular cerebral com deficits cognitivos	430.00-432.90; 433.01; 433.11; 433.21; 433.31; 433.81; 433.91; 434.00-434.91; 436.00; 437.50; 437.60; 438.10-438.19 e 438.00
3.2	Acidente cerebrovascular e envolvimento dos nervos cranianos	S3.1 + 377.49; 378.50-378.56; 350.80-352.90; 781.10; 388.50; 389.10-389.20; 389.80; 389.90; 787.20; 369.00-369.18; 369.21; 369.22; 369.24; 369.40-369.69 e 369.71-369.73 438.82
3.3	Acidente cerebrovascular com sinais cerebelares ou monoplegia	S3.1-S3.2 + 334.30; 379.50; 379.52; 379.53; 379.55-379.59; 781.30 e 781.00 438.30-438.52
3.4	Acidente cerebrovascular com hemiparesia ou hemiplegia	S3.1-S3.3 + 342.00-342.92 438.20-438.22
3.5	Acidente cerebrovascular com tetraplegia ou síndrome parálitico-bilateral	S3.1-S3.4 + 344.00-344.09 438.53
3.6	Acidente cerebrovascular e coma	S3.1-S3.4 + 780.00 e 780.01
4.0	Acidente cerebrovascular e morte	

Fonte: Gonnella, Louis e Gozum, 1999

Disease Staging

Grandes Agrupamentos de Doenças (GADs)

GADs	Grandes Agrupamentos de Doenças
BR	Doenças da mama
CN	Doenças do sistema nervoso central
CV	Doenças cardiovasculares e do coração
DE	Doenças dermatológicas
EN	Doenças do ouvido, nariz e garganta
GI	Doenças gastrointestinais
GY	Doenças ginecológicas
HB	Doenças hepatobiliares
HE	Doenças hematológicas
IM	Doenças Imunológicas
LY	Doenças linfáticas
MG	Doenças dos órgãos genitais masculinos
MS	Doenças musculoesqueléticas
ND	Doenças endócrinas e metabólicas
OB	Doenças obstétricas
OP	Doenças oftalmológicas
PN	Doenças do sistema nervoso periférico
PS	Doenças psiquiátricas
RN	Doenças renais
RS	Doenças respiratórias
SY	Doenças multisistémicas
VS	Doenças vasculares (excepto coração)
NN	Doenças não especificadas
OT	Outras doenças

ANEXO II – DIAGNOSIS RELATED GROUPS

- Portaria nº 189/2001 de 9 de Março
- Grandes Categorias de Diagnóstico (GCDs)

Diagnosis Related Groups

Portaria nº 189/2001 de 9 de Março

Ver CD – directoria Anexo II

Diagnosis Related Groups

Grandes Categorías de Diagnóstico (GCDs)

GCDs	Grandes Categorias Diagnosticas
GCD0	Pré-grandes categorias diagnósticas
GCD1	Doenças e perturbações do sistema nervoso
GCD2	Doenças e perturbações do olho
GCD3	Doenças e perturbações do ouvido, nariz, boca e garganta
GCD4	Doenças e perturbações do aparelho respiratório
GCD5	Doenças e perturbações do aparelho circulatório
GCD6	Doenças e perturbações do aparelho digestivo
GCD7	Doenças e perturbações do sistema hepatobiliar e pâncreas
GCD8	Doenças e perturbações do sistema musculoesquelético e tecido conjuntivo
GCD9	Doenças e perturbações da pele, tecido celular subcutâneo e mama
GCD10	Doenças e perturbações endócrinas, nutricionais e metabólicas
GCD11	Doenças e perturbações do rim e do aparelho urinário
GCD12	Doenças e perturbações do aparelho genital masculino
GCD13	Doenças e perturbações do aparelho genital feminino
GCD14	Gravidez, parto e puerpério
GCD15	Recém-nascidos e lactentes com afecções do período perinatal
GCD16	Doenças e perturbações do sangue/órgãos hematopoiéticos e doenças imunológicas
GCD17	Doenças e perturbações mieloproliferativas e mal diferenciadas
GCD18	Doenças infecciosas e parasitárias (sistémicas ou de localização não específica)
GCD19	Doenças e perturbações mentais
GCD20	Uso de álcool/droga e perturbações mentais orgânicas induzidas por álcool ou droga
GCD21	Traumatismos, intoxicações e efeitos tóxicos de drogas
GCD22	Queimaduras
GCD23	Factores com influência no estado de saúde e outros contactos com serviços de saúde
GCD24	Traumatismos múltiplos significativos
GCD25	Infecções pelo vírus da imunodeficiência humana
GCD99	Grupos com procedimentos no bloco operatório não relacionados com o diagnóstico principal

Fonte: Portaria nº 189/2001 de 9 de Março

ANEXO III – DEMORA MÉDIA - OUTLIERS

- Total de Doentes Saídos, Média, Mediana, 1º e 3º Quartis, Distância Quartal e as Barreiras de “Outliers” Inferiores e Superiores por DRG
- Doentes Saídos, Outliers, Inliers e % de Outliers por DRG
- Doentes Saídos e Demora Média por DRG, Total da População e Inliers e Demora Média Esperada Inliers
- % Outliers por Hospital; Reduções na Demora Média por Hospital
- Demora Média Observada e Esperada Original e Esperada Recalibrada por Hospital

Ver CD – directoria Anexo III

ANEXO IV – ÍNDICES DE CASEMIX

- Total da Produção por Hospital
- Casos Cirúrgicos por Hospital
- Casos Médicos por Hospital
- Grandes Agrupamentos de Doenças (GADs)
- Índice de Casemix Composto por Hospital

Ver CD – directoria Anexo IV

ANEXO V – CONCENTRAÇÃO DA PRODUÇÃO

- % de Produtos Diferentes por Hospital de acordo com os DRGs e com o Disease Staging
- % de Produtos Diferentes Responsáveis por 80% da Produção por Hospital, de acordo com os DRGs e com o Disease Staging

Ver CD – directoria Anexo V

ANEXO VI – COMPLEXIDADE DOS CASOS POR HOSPITAL

- Total da Produção
- Casos Cirúrgicos
- Casos Médicos

Ver CD – directoria Anexo VI

ANEXO VII – GRAVIDADE DOS CASOS POR HOSPITAL

- Total da Produção
- Casos Cirúrgicos
- Casos Médicos

Ver CD – directoria Anexo VII

ANEXO VIII – COMPLEXIDADE E GRAVIDADE DOS CASOS POR HOSPITAL

- Total da Produção
- Casos Cirúrgicos
- Casos Médicos

Ver CD – directoria Anexo VIII

ANEXO IX – COMPLEXIDADE E GRAVIDADE DOS CASOS POR DOENÇA (GADS)

- Total da Produção
- Casos Cirúrgicos
- Casos Médicos

Ver CD – directoria Anexo IX

ANEXO X – EFECTIVIDADE

- Taxa de Mortalidade por Hospital – Total da Produção, Casos Cirúrgicos e Casos Médicos
- Taxa de Mortalidade por Doença (GADs)
- Taxa de Mortalidade Esperada por Hospital – Total da Produção, Casos Cirúrgicos e Casos Médicos
- Taxa de Mortalidade Padronizada por Hospital
- Ordenação dos Hospitais por Taxa de Mortalidade Observada e por Taxa de Mortalidade Padronizada
- Efectividade por Hospital (Z score) – Total da Produção, Casos Cirúrgicos e Casos Médicos

VER CD – directoria Anexo X

ANEXO XI – EFICIÊNCIA

- Demora Média por Hospital
- Demora Média Observada e Esperada por Estadio e por Hospital
- Eficiência por Hospital – Métodos Directo e Indirecto
- Eficiência por Hospital (Z score) – Total da Produção, Casos Cirúrgicos e Casos Médicos – Ordenação

Ver CD – directoria Anexo XI

ANEXO XII – DESEMPENHO GLOBAL

- Ordenação dos Hospitais – Método Indirecto
- Ordenação dos Hospitais – Método Directo
- Comparação com Efectividade e Eficiência – Total da Produção, Casos Cirúrgicos e Casos Médicos

Ver CD – directoria Anexo XII

